

Centro de Investigación Científica y de
Educación Superior de Ensenada

PAQUETIZACION DE VIDEO Y SU TRANSMISION
EN REDES DE COMUNICACIONES
MULTISERVICIOS

TESIS
MAESTRIA EN CIENCIAS

MIGUEL ANGEL COTA MIRANDA

Ensenada, Baja California, Mexico. Octubre de 1995.

RESUMEN de la tesis de Miguel Angel Cota Miranda, presentada como requisito parcial para la obtención del grado de MAESTRO en CIENCIAS en ELECTRÓNICA Y TELECOMUNICACIONES. Ensenada, Baja California, México, Octubre de 1995.

PAQUETIZACIÓN DE VIDEO Y SU TRANSMISIÓN EN REDES DE COMUNICACIONES MULTISERVICIOS

Resumen aprobado por:

Con la aparición de nuevas tecnologías de redes de conmutación de paquetes existen alternativas para la transmisión de video digital comprimido que comunmente ha sido realizada mediante redes de conmutación de circuitos. Las redes de conmutación de paquetes ofrecen ventajas con respecto a los sistemas de conmutación de circuitos como calidad constante del servicio y mayor eficiencia en la utilización del ancho de banda. Sin embargo, las redes de conmutación de paquetes generalmente producen retardos variables y posibles pérdidas de paquetes que pueden afectar la calidad del servicio.

En el presente trabajo son de particular importancia las condiciones de operación que debe de observar una red de conmutación de paquetes para asegurar la transmisión exitosa del servicio de videoconferencia. Se estudian los estándares más adecuados para la compresión de las señales de video, las condiciones de interactividad requeridas por el servicio, los retardos extremo a extremo desde el proceso de codificación en el extremo transmisor, la transmisión sobre la red de conmutación de paquetes y la decodificación en el extremo receptor. Se analizan los casos para la codificación a tasa de bit constante y tasa de bit variable, obteniendo las ventajas y desventajas de cada uno de estos. Para obtener las condiciones de operación óptimas se modela la tasa de bit variable mediante un modelo de Markov de nacimiento y muerte, y este modelo es empleado en el análisis del sistema de almacenamiento y envío del nodo de acceso para estimar la distribución de la probabilidad de pérdidas de paquetes en el almacenador temporal.

Como resultado de este estudio se obtienen los parámetros de operación óptimos tales como el retardo de transmisión máximo permisible por el servicio, el tamaño de almacenador y la tasa del canal de alta velocidad.

ABSTRACT of the thesis, presented by Miguel Angel Cota Miranda, in order to obtain the **MASTER of SCIENCE DEGREE in TELECOMMUNICATIONS AND ELECTRONICS**. Ensenada, Baja California, Mexico, October 1995.

PACKET VIDEO AND ITS TRANSMISSION OVER MULTISERVICES COMMUNICATIONS NETWORKS.

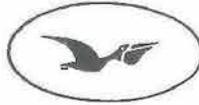
Approved by:

With the recent technologies in packet switching networks, there are choices for transmit compressed digital video than usually has been made through circuit switching networks. Indeed, packet switching networks offers advantages with respect to the circuit switching systems like high quality services and best efficiency on bandwidth use. Nevertheless, packet switching networks generally produces no constant delay and possible packet loss than could affect the quality on service.

In this work, are very important the packet switching networks operation's conditions in order to guarantee the videoconference service. We studied the more used standards in video compressing systems, the end to end delays since the coding process in the transmitting end, the transmission over the packet switching network, and the decoding process in the receiver end. We analyze, the cases for constant bit rate and variable bit rate, obtaining its advantages and disadvantage. To obtain the optimum conditions, the variable bit rate is modeled using a birth and death Markov Process, which is used in the analysis of the storage and forward system in order to obtain the packet loss in the buffer.

As an important result of this work we obtain the optimum networks's operative values like the maximum permissible transmission delay by the service, the buffer's length and the rate of the high speed channel.

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE
EDUCACIÓN SUPERIOR DE ENSENADA



CICESE

DIVISIÓN DE FÍSICA APLICADA

DEPARTAMENTO DE ELECTRÓNICA Y TELECOMUNICACIONES

**PAQUETIZACIÓN DE VIDEO Y SU TRANSMISIÓN EN REDES DE
COMUNICACIONES MULTISERVICIOS**

TESIS

que para cubrir los requisitos necesarios para obtener el
grado de **MAESTRO EN CIENCIAS** presenta:

MIGUEL ANGEL COTA MIRANDA

Ensenada Baja California, México. Octubre de 1995.

DEDICATORIA

El presente trabajo lo dedico con mi más profundo cariño y respeto:

A la memoria de mi padre Roberto Pablo Cota Moreno,

que al igual que en las demás etapas de mi vida, impulsaste con gran ánimo el inicio de esta que termina, y a pesar de que ya no estás físicamente entre nosotros, tu recuerdo, tus consejos, tu sentido de la honradez, y tu enorme voluntad hacia el trabajo han sido los mayores impulsos en el desarrollo y culminación de la presente etapa.

A mi madre Martha Miranda Amador,

que con abnegación y respeto a mis deseos de superación, haz permitido mi ausencia en el hogar que con tanto amor han forjado.

Con todo mi cariño:

a mis hermanos Nieves, Roberto, Alma Rosa, Carlos, Alejandro, Martha, Mario, Cecilia, Leonor, Jorge, Leticia, Francisco y Verónica,

Por su invaluable apoyo y muestras de afecto, por su gran espíritu fraternal en los momentos difíciles, su constante motivación a que obtuviera más y mejores logros, porque juntos hemos formado una gran familia: la dinastía Cota Miranda.

Con mucho cariño:

a mis cuñados, sobrinos y demás familiares

que a pesar de la lejanía, estuvieron constantemente brindándome su apoyo.

AGRADECIMIENTOS

A Dios, por brindarme siempre su luz divina en los momentos más difíciles, y por hacerla más brillante a medida que más desatendía de ella.

A mi amigo y director de tesis M.C. David Covarrubias Rosales, por su invaluable apoyo, confianza y amistad brindados, los cuales fueron factores fundamentales en el desarrollo de este trabajo.

A los miembros del comité de tesis Dr. Francisco Javier Mendieta Jiménez, Dr. Ciro Andrés Martínez García-Moreno y M.C. Roberto Conte Galván, por su atinada orientación y la revisión del manuscrito.

A mi amigo M.C. Jorge Preciado Velasco, por su amistad y confianza durante mi estancia en el CICESE.

A los compañeros y amigos de generaciones anteriores Zarina Guillén, Ana Maria Terpán, Salvador Durazo, Ramón Soto, Julio Rolón, Carlos Brizuela, Abel Rodríguez, Jorge Peña, Carlos Merla, Miguel Gómez, Iván Nieto, Alejandro Márquez, L. A. Cetto y muy especialmente a Manuel Corona, Horacio Martínez, Francisco Rangel y Hernando Silva por la motivación brindada.

A los compañeros de generación Alvaro Ramírez, Calixto Estrada, Alejandro Cervantes, Juan Carlos Murrieta, Oscar Iván Lepe y Luis Castro, más que un grupo, formamos una auténtica familia: la Familia Capullo.

A mi amigo y tocayo Miguel Martínez, por su amistad incondicional.

A los ópticos Veneranda, Pablo Loza, Victor Coello, y Rodolfo, por su amistad y compañerismo.

A mi paisana Zarina Talamantes, que comparte el orgullo de haber nacido en nuestra entrañable tierra sudcaliforniana, gracias por tu amistad.

AGRADECIMIENTOS (continuación)

A los compañeros de nuevas generaciones, especialmente a Alvaro López, Alberto Alvarado y Mario López , y a los demás que por falta de espacio y premura de tiempo no me es posible mencionar, les agradezco su amistad.

Por supuesto, a “las muchachas” Rosa Chávez, Rosa Martha López, Aidel López y Maripaz Luna, por su amistad, la cual espero que permanezca como una piedrita de diamante: para siempre.

Al CONACyt por su apoyo económico.

Al CICESE por impulsar mi superación.

CONTENIDO

Página

I INTRODUCCIÓN

I.1 Antecedentes	1
I.2 Motivaciones del trabajo	3
I.3 Objetivo	4
I.4 Metas	4
I.5 Alcances	5
I.6 Infraestructura empleada	6
I.7 Organización del trabajo	6

II CARACTERÍSTICAS DE LA SEÑAL DE VIDEO

II.1 Formato original	9
II.1.1 Contenido de información del servicio	10
II.1.2 Resolución espacial	13
II.1.3 Resolución temporal	14
II.1.4 Espacio de colores	15
II.2 Requerimientos de ancho de banda para video digital	16
II.3 Compresión de video	18
II.3.1 Técnicas de compresión de imágenes	19
II.3.1.1 Redundancia de código	20
II.3.1.2 Redundancia entre elementos de imagen	21
II.3.1.3 Redundancia psicovisual	22
II.3.2 Modelos de compresión de imágenes	23
II.3.2.1 Codificador y decodificador fuente	24
II.3.2.2 Codificador y decodificador de canal	26
II.3.3 Estándares para compresión de video	27
II.3.4 Características de la tasa de bit para video comprimido	28

III DESARROLLO EXPERIMENTAL DE LA COMPRESIÓN DE VIDEO

III.1 Esquema experimental de trabajo	31
III.2 Experimentos a tasa de bit constante	33
III.3 Experimentos a tasa de bit variable	37

IV ANÁLISIS DE RETARDOS PARA VIDEO PAQUETIZADO

IV.1 Identificación de retardos extremo a extremo para video paquetizado	47
IV.1.1 Retardos de codificación y decodificación de la señal de video	50
IV.1.1.1 Codificación a tasa de bit constante	51
IV.1.1.2 Codificación a tasa de bit variable	58

CONTENIDO (continuación)

Página

IV.1.2 Retardo dentro de la red de conmutación de paquetes	61
--	----

V CARACTERIZACIÓN DE LA TASA DE BIT PARA VIDEO PAQUETIZADO

V.1 Caracterización para la codificación de video a tasa de bit constante	64
V.2 Caracterización para la codificación a tasa de bit variable	65
V.3 Modelaje para la codificación de videoconferencia a tasa de bit variable	69
V.3.1 Caracterización de una secuencia pseudoaleatoria	69
V.3.2 Modelo para la tasa de bit variable	75

VI ANÁLISIS DE RETARDOS EN LA RED DE CONMUTACIÓN DE PAQUETES

VI.1 Retardo generado por la paquetización	86
VI.1.1 Fuente de video con tasa de bit constante	87
VI.1.2 Fuente de video a tasa de bit variable	90
VI.2 Análisis de retardos de almacenamiento temporal	94
VI.2.1 Transmisión sin retardo de almacenamiento	96
VI.2.2 Transmisión con retardo de almacenamiento	99
VI.3 Análisis de resultados	104

VII CONCLUSIONES Y RECOMENDACIONES

VII.1 Conclusiones	115
VII.2 Aportaciones	121
VII.3 Recomendaciones	122

LITERATURA CITADA	123
--------------------------	-----

APÉNDICE A: Descripción del estándar de compresión ITU-T H.261	125
---	-----

LISTA DE FIGURAS

Figura	<u>Página</u>
1. Proceso de la transmisión de video paquetizado.	2
2. Parámetros que caracterizan a una secuencia de video.	10
3. Imagen típica de videotelefonía.	12
4. Imágenes típicas de videoconferencia (a) Secuencia "Miss America" (b) Secuencia "Salesman".	12
5. Etapas para un sistema general de compresión de imágenes.	23
6. Procesos de (a) codificación y (b) decodificación fuente.	24
7. Tasa de bit y calidad generada por los sistemas de compresión de video: (a) tasa de bit fijo y (b) tasa de bit variable.	29
8. Proceso experimental para la compresión de video.	32
9. Efecto de la codificación a tasa de bit constante sobre la calidad de las imágenes. (a) 64 kbps, (b) 128 kbps, (c) 380 kbps, (d) 1.5 Mbps.	35-36
10. Reducción en la cantidad de datos comprimidos al modificar el valor de cuantización en un esquema de compresión de transformada de coseno discreto combinado con codificación de longitud variable.	38
11. Imágenes decodificadas después de ser comprimidas con diversos valores de cuantización. (a) $q=1$, (b) $q=10$, (c) $q=20$, (d) $q=30$.	40-41
12. Tasa de bit generada por la secuencia "Salesman".	42
13. Elementos en la transmisión de video paquetizado en los cuales se producen retardos constantes o variables.	48
14. Diagrama de temporización del flujo de video.	49
15. Fuentes posibles de retardos en el extremo transmisor de la señal de video.	52
16. Fuentes de retardos en el extremo receptor de la señal de video.	55
17. Histograma del proceso de bits codificados por cuadro.	66
18. Función de autocovarianza para el proceso de bits codificados por cuadro.	67
19. Secuencia pseudoaleatoria generada para analisis comparativo.	70
20. Histograma para la secuencia pseudoaleatoria.	71
21. Autocovarianza de la secuencia pseudoaleatoria.	72
22. Secuencia pseudoaleatoria para 10000 datos.	73
23. Histograma para el caso de 10000 números pseudoaleatorios.	73
24. Función de covarianza para el caso de 10000 números pseudoaleatorios.	74
25. Diagrama de transición de estados para el modelo empleado	76
26. Probabilidades de la tasa de bit total del modelo, parametrizada por M.	80
27. Autocovarianza de la tasa de bit total. $N=1$.	82
28. Distribuciones de probabilidad para diversos valores de σ .	84
29. Retardo producido por el proceso de formación de la celda a tasa de bit constante.	88

LISTA DE FIGURAS (continuación)

<u>Figura</u>		<u>Página</u>
30.	Variación del retardo de paquetización en función de la tasa de bit.	89
31.	Tasa de bit de la secuencia "salesman" expresada en bits por segundo.	91
32.	Retardo de paquetización a tasa variable para la secuencia "salesman".	92
33.	Modelo para el nodo de acceso.	94
34.	Tasa de servicio c mínima para no producir retardos de almacenamiento.	98
35.	Comportamiento del modelo para diversos valores de M . $N=1$, $\rho=0.9996$, $c=720288$ bps.	104
36.	Comportamiento de las pérdidas para el caso de múltiples fuentes. $\rho=0.9996$.	107
37.	Requerimientos en el tamaño del almacenador en función del número de fuentes multicanalizadas y de la probabilidad de pérdidas que genere.	109
38.	Comportamiento de las pérdidas en función del número de fuentes y con un determinado tamaño de almacenador.	110

LISTA DE TABLAS

<u>Tabla</u>		<u>Página</u>
I.	Tolerancia de los servicios de video a movimientos de cámara y a los cambios de escena.	11
II.	Resolución espacial para diferentes calidades de video.	14
III.	Resolución temporal para diferentes estándares de video.	15
IV.	Anchos de banda para diversos formatos de video.	17
V.	Valores numéricos de los parámetros operativos para videoconferencia.	113

PAQUETIZACIÓN DE VIDEO Y SU TRANSMISIÓN EN REDES DE COMUNICACIONES MULTISERVICIOS

I INTRODUCCIÓN

I.1 Antecedentes

Las redes de conmutación de paquetes han sido utilizadas eficazmente alrededor del mundo y en México para la transmisión de servicios no sensibles a retardos tales como los datos, por lo que su infraestructura y protocolos están orientados hacia sus requerimientos. Sin embargo, nuevas necesidades en cuanto al tipo de servicios han surgido por parte de los usuarios de estas redes, en particular la de transmitir servicios sensibles a retardos como lo son voz, audio y video.

De especial interés es el servicio de video digital comprimido, el cual ha sido transmitido mediante redes basadas en conmutación de circuitos. Sin embargo, la adaptación de los codificadores de video a tales redes puede resultar en la degradación del servicio, lo que resulta indeseable. Con las nuevas tecnologías de redes rápidas de conmutación de paquetes, la degradación del servicio puede reducirse notablemente, por lo que se espera que los servicios de video ocupen una fracción importante del ancho de banda ofrecido por estas nuevas generaciones de redes de paquetes [Karlsson y Vetterli, 1989].

En la figura 1 se describe el proceso de la transmisión de video paquetizado [Ohta, 1994]. Una secuencia de video es digitalizada, codificada y comprimida mediante el codificador. A la salida del codificador se obtiene información con un determinado

comportamiento en el tiempo, al cual se le denomina tasa de bit (en inglés nombrado bit rate). Tal información es multicanalizada con otros servicios de video o de naturaleza diferente (datos, voz) en un punto de la red conocido como nodo de acceso, y convertida a las unidades básicas de transmisión (paquetes) asociadas a la red de alta velocidad, los cuales son enviados mediante una cierta trayectoria o circuito de comunicación dentro de la red. Al otro extremo de la red, los paquetes son reensamblados y la información es descomprimida y vuelta a su formato original en el decodificador.

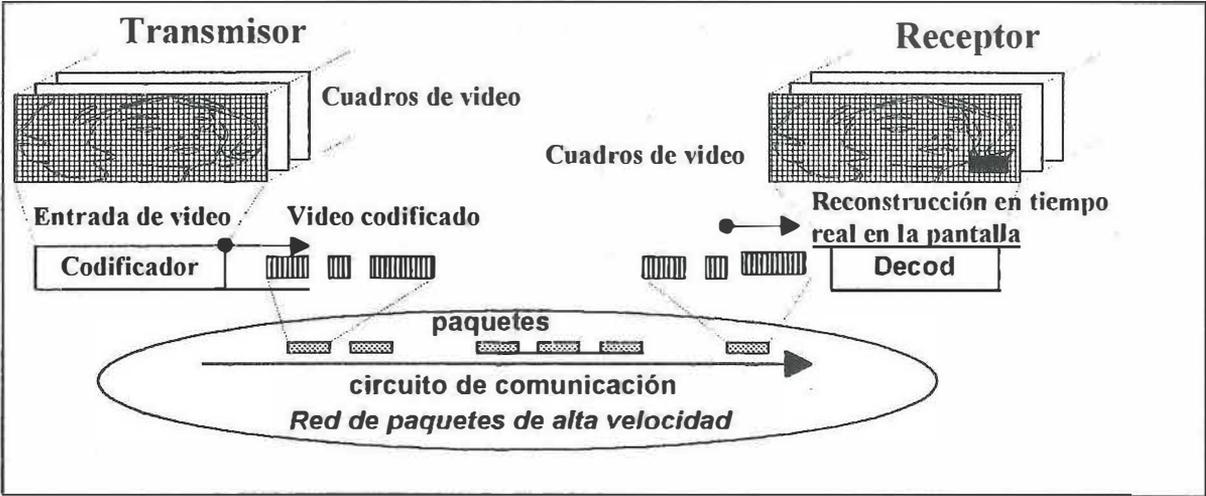


Figura 1. Proceso de la transmisión de video paquetizado

Sin embargo, la red de paquetes no garantiza la transmisión exitosa de todos los paquetes, por lo que algunos de estos pudieran extraviarse, o ser entregados demasiado tarde como para ser utilizados a tiempo en la reconstrucción de la imagen en el receptor,

provocando una degradación en la calidad del servicio. Por lo tanto, es necesario estudiar las condiciones de operación de la red de alta velocidad que garantice, si no la entrega completa de todos los paquetes, sí un número limitado de paquetes perdidos y/o retrasados para evitar que la calidad del servicio se deteriore.

Este trabajo estudia las características de la señal de video, en especial las obtenidas de servicios de videoconferencia, en cuanto a su generación, codificación y compresión, tolerancia a retardos de transmisión, los diferentes procesos a los cuales se somete durante la transmisión en una red de paquetes y sus posibles efectos en la calidad del servicio. En base a este análisis, se desea obtener las condiciones de operación antes mencionadas.

I.2 Motivaciones del trabajo

Existen diferentes motivos que dan pie a la realización de este trabajo, entre los más importantes se pueden mencionar:

- La transmisión de video mediante redes rápidas de conmutación de paquetes ofrece una serie de ventajas con respecto a la utilización de conmutación de circuitos, como calidad del servicio constante y utilización más eficiente del ancho de banda [Karlsson y Vetterli, 1989].
- En México se ha presentado un aumento considerable en el uso de redes de conmutación de paquetes, sobre todo para transmitir servicios no sensibles al tiempo como los datos. Recientemente se han instalado en las principales ciudades del país, redes rápidas de conmutación de paquetes las cuales ya soportan servicios sensibles al tiempo como la voz. La tendencia es la integración de todos los servicios en una sola red, por lo que la inclusión

de los servicios visuales es la siguiente meta en tal dirección. Sin embargo, existen pocos estudios dedicados a abordar la problemática asociada a la transmisión de video paquetizado, y este trabajo es un paso en ese sentido.

-Es importante determinar bajo qué condiciones una red rápida de conmutación de paquetes puede transmitir el servicio de video sin degradar la calidad.

-Es un tema de actualidad ya que implica abordar el estado del arte de las redes de comunicaciones.

I.3 Objetivo

Modelaje y simulación para la optimización de la transmisión de la señal de videoconferencia en redes rápidas de conmutación de paquetes .

I.4 Metas

Para alcanzar el objetivo anterior se plantean las siguientes metas u objetivos particulares:

- I. Estudio sobre las características básicas de la señal de video, en cuanto a su generación, codificación y compresión.
- II. Análisis de retardos extremo a extremo en la transmisión de video paquetizado.
- III. Caracterización y modelaje de la tasa de bit generada por secuencias típicas de videoconferencia.

IV. Modelaje del nodo de acceso a la red de conmutación de paquetes, para obtener la distribución del retardo en el almacenador temporal y la correspondiente probabilidad de pérdidas de datos.

V. Estimación de valores óptimos para los diferentes parámetros asociados a la red que permitan la transmisión del servicio de videoconferencia.

I.5 Alcances

A partir del objetivo y metas planteadas anteriormente, los alcances del presente trabajo se pueden acotar a lo siguiente:

I. Análisis de los retardos generados en la transmisión de video en redes de paquetes.

II. Análisis y modelaje del proceso de almacenamiento temporal en el nodo de acceso a la red de paquetes, que permitan concluir acerca de las condiciones de operación más favorables para la transmisión de video.

III. Se analizan y modelan únicamente las tasas de bit generadas al codificar secuencias típicas de videoconferencia y sin cambios de escena.

IV. El análisis de procesos adicionales que pueden afectar la calidad de la señal paquetizada, tales como la sincronización de las terminales de video, los esquemas de tratamiento de errores más apropiados, paquetes perdidos y/o retardados, se encuentran fuera del alcance de este trabajo [Karlsson y Vetterli, 1989; Ohta, 1994].

I.6 Infraestructura empleada

Para llevar a cabo todo el desarrollo experimental involucrado en este trabajo, la infraestructura empleada fue:

- Estaciones de trabajo SUN SPARC WORKSTATION del Laboratorio de Procesamiento Digital de Señales.
- Secuencias de video digitalizadas obtenidas de Internet.
- Simulador para codificación y compresión de video p64 para el estándar ITU-T H.261, obtenido de Internet.
- Programa de aplicación Mathcad 5.0 y ambiente de programación Turbo C++.
- Computadora personal AT486

I.7 Organización del trabajo

En el capítulo II se muestran las características de la señal de video, los requerimientos en su codificación a formato digital, y la necesidad de comprimir el volumen de datos generados por el proceso de digitalización. Se describen las técnicas de compresión de imágenes más utilizadas en la actualidad y los estándares de codificación y compresión de video y su influencia en la tasa de bit generada. Se elige al esquema H.261 para ser utilizado en la parte experimental del trabajo.

En el capítulo III se presenta el desarrollo experimental para el proceso de codificación y compresión de video utilizando un simulador para el estándar H.261. Una cuestión importante es observar la degradación subjetiva generada por este esquema de compresión

de video. También se realiza un experimento con una versión modificada del esquema H.261 que genere tasa de bit variable con la finalidad de observar el resultado en la calidad subjetiva del servicio y obtener datos reales para ser utilizados en el análisis de transmisión de video a tasa de bit variable. Para los experimentos se utilizan secuencias de videoconferencia típicas.

En el capítulo IV se realiza un análisis para obtener los retardos extremo a extremo en la transmisión de video paquetizado, identificando a los procesos que producen retardos variables. Se presentan dos casos: el primero para un codificador a tasa de bit constante (H.261), y después para un esquema simplificado de H.261 (para fines de análisis) que produce tasa de bit variable. Son obtenidos los intervalos de operación que no producen degradaciones al servicio por causa de retardos excesivos. La importancia que tiene el retardo presente en el almacenador temporal (cola de espera) del nodo de acceso, es de particular interés en este análisis.

Para conocer la distribución del retardo en el almacenador temporal del nodo de acceso, se requiere caracterizar a la información que esté arribando al mismo, i.e., conocer el comportamiento estadístico de la tasa de bit generada al codificar y comprimir video; y obtener un modelo para este proceso. Esto se lleva a cabo en el capítulo V. Los datos a analizar son los presentados en el capítulo III.

Una vez modelado el proceso de la tasa de bit generada por los codificadores de video, el capítulo VI aborda el problema de modelar el nodo de acceso a la red, con la finalidad de conocer la distribución del retardo generado en el almacenador temporal. También se desea

conocer si se pueden producir pérdidas por sobreflujo del almacenador y cual sería el comportamiento de éste en el caso de que ocurrieran.

Para finalizar, en el capítulo VII se presentan las conclusiones generales del trabajo, las aportaciones y recomendaciones.

II CARACTERÍSTICAS DE LA SEÑAL DE VIDEO

La transmisión de video paquetizado involucra la digitalización de la señal de video que comúnmente es generada por las cámaras de video en forma analógica. En este capítulo se revisan las características más importantes de la señal de video para el proceso de conversión a formato digital; también los diferentes tipos de servicio existentes, dando un especial énfasis a videoconferencia. También se obtienen los anchos de banda mínimos requeridos para transmitir los servicios de video digital y, como una consecuencia, se concluye en la necesidad de reducir la cantidad de datos generados por la digitalización de las señales de video. De manera muy breve se presentan las técnicas de compresión más ampliamente utilizadas en los sistemas de codificación de video estandarizados actualmente y su influencia en el ancho de banda requerido y en la calidad subjetiva del servicio.

Lo anterior permitirá seleccionar el sistema de codificación y compresión de video que sirva como referencia al desarrollo experimental del presente trabajo.

II.1 Formato original

El video se considera básicamente como una sucesión de imágenes desplegadas en una pantalla con una determinada frecuencia de repetición. Para los fines de este trabajo, una secuencia de video se puede caracterizar por cuatro parámetros: el contenido de información del servicio, el espacio de colores, la resolución espacial y la resolución temporal, como se muestra en la figura 2. A continuación se describen brevemente cada uno de estos parámetros.

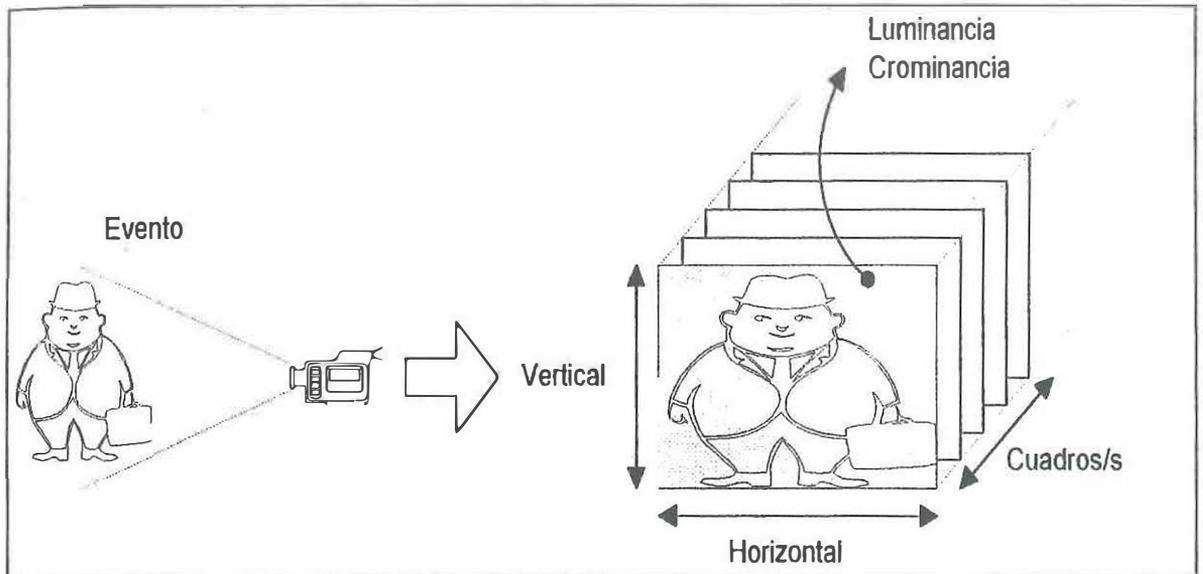


Figura 2. Parámetros que caracterizan a una secuencia de video.

II.1.1 Contenido de información del servicio

El contenido de información de una secuencia de video depende de la aplicación a la que se destina la transmisión. Existe una diversidad de eventos que pudieran servir como fuentes de información para las señales de video, pero éstas básicamente se han agrupado en tres categorías o calidades de servicio [Ohta, 1994]: videotelefonía, videoconferencia y televisión comercial. La diferencia entre éstos radica en los movimientos de objetos o personas dentro de la secuencia, movimientos de la cámara de video y cambios de escena.

En la tabla I se muestra la tolerancia típica de estos servicios a las dos últimas características.

Tabla I. Tolerancia de los servicios de video a movimientos de cámara y a los cambios de escena.

Tipo de servicio	Movimientos de cámara	Cambios de escena
Videotelefonía	No existen	No existen
Videoconferencia	Generalmente no son necesarios	No existen
TV comercial	Bastante frecuentes	Bastante frecuentes

Para ilustrar las características de las imágenes y los movimientos típicos de personas y objetos en los servicios de videotelefonía y videoconferencia, a continuación se muestran una serie de imágenes obtenidas de secuencias de video que se han utilizado como referencias en diversas investigaciones realizadas para determinar la eficiencia de los diferentes sistemas de compresión de imágenes y de video. En la figura 3, se presenta una imagen típica de videotelefonía, obtenida de la secuencia denominada ‘Suzie’. Como puede observarse, únicamente aparecen cabeza, hombros y la mano que sostiene el auricular. El fondo de la imagen permanece fijo. Los movimientos que pudieran esperarse de las personas en este tipo de servicio son de manos, brazos, labios y de la cabeza en general.

La figura 4 presenta dos imágenes típicas de videoconferencia. La figura 4-a pertenece a la secuencia ‘Miss America’, mientras que la figura 4-b se obtiene de la secuencia ‘Salesman’. Las personas están sentadas frente a un escritorio, por lo que pueden observarse de los hombros hacia arriba. De esta manera, los movimientos que pueden esperarse de los sujetos en este tipo de servicio es de brazos, cabeza, labios. El fondo de la



Figura 3. Imagen típica de videotelefonía



(a)



(b)

Figura 4. Imágenes típicas de videoconferencia: (a)Secuencia “Miss America”, (b)Secuencia “Salesman”.

escena permanece fijo y puede ocupar una buena parte de la imagen, como se observa en estos casos.

Para las secuencias de TV comercial no se presentan imágenes, dado que en este servicio se permite toda la diversidad posible en el contenido de información (películas, deportes, noticiarios, conferencias, escenarios naturales, dibujos animados, etcétera), sin limitación en cuanto a movimientos de sujetos u objetos animados, movimientos de cámara, características del fondo de escena, cambios de escena, etcétera; por lo que este servicio contiene mucha mayor información que los presentados anteriormente.

Como ya se ha mencionado, este trabajo está dirigido al análisis de servicios de videoconferencia. Por lo tanto, se considera que la secuencia de videoconferencia no presenta movimientos de cámara ni cambios de escena.

II.1.2 Resolución espacial

Una imagen o cuadro se puede considerar que está subdividido en un agrupamiento de pequeños elementos que se conocen como píxeles. Si se aumentan la cantidad de datos que conforman la imagen, entonces la nitidez de ésta tenderá a mejorar. Sin embargo, aumentar arbitrariamente la cantidad de píxeles en una imagen conlleva dos problemas: el primero es que se requieren cámaras de video más sofisticadas y caras, el segundo, como se verá en secciones posteriores, el ancho de banda mínimo de transmisión puede crecer a valores inconvenientes de manejar de acuerdo a las capacidades del canal que se utilice.

La resolución espacial se representa por dos valores que son la resolución horizontal y la resolución vertical, como se observa en la figura 2. Es decir, si se considera a una imagen como un rectángulo, entonces la resolución horizontal sería la cantidad de elementos de imagen definidos en dirección horizontal y la resolución vertical estaría dada por el número de elementos presentes en una línea en dirección vertical. La resolución espacial para diferentes tipos de servicios se muestra en la tabla II [Ohta, 1994].

Tabla II. Resolución espacial para diferentes calidades de video

Tipo de imagen	Resolución espacial
Videotelefonía	176x144 (QCIF-H.261)
Videoconferencia	352x288 (CIF) ó QCIF
TV digital	720x480 (NTSC)
HDTV digital	1280x720 (propuesta americana)

Como se ha venido destacando, en el presente trabajo se consideran secuencias de videoconferencia. En particular, la resolución espacial utilizadas en la parte experimental es QCIF-176x144, definida en el estándar ITU-T H.261.

II.1.3 Resolución temporal

La resolución temporal se refiere a la cantidad de cuadros que se transmiten por unidad de tiempo, con la finalidad de dar la sensación de movimiento al observador humano (figura 2). Dependiendo del estándar de video, se definen diferentes valores para la resolución

temporal. La tabla III muestra la frecuencia de repetición de cuadros utilizados por los estándares de video más comunes en la actualidad.

Tabla III. Resolución temporal para diferentes estándares de video

Estándar de video	Resolución temporal (cuadros/s)
NTSC	30
PAL	25
H.261	Variable (máximo 30)

El estándar H.261 permite variar la resolución temporal hasta 30 cuadros/s, con la finalidad de elegir la calidad deseada del servicio. En la parte experimental del trabajo se utilizó la máxima resolución temporal posible, es decir, 30 cuadros/s.

II.1.4 Espacio de colores

Si se intenta transmitir video monocromático, la información de cada pixel que forma una imagen se representa mediante un valor proporcional a su nivel de brillantez, denominado luminancia.

Para video a color, se requiere información adicional que permita sintetizar, aparte del nivel de brillantez de cada elemento, su color o crominancia (figura 2).

Cada estándar de video define su propio espacio de colores. En este trabajo, se utilizó el espacio de colores definido por el estándar H.261, conocido como YUV.

II.2 Requerimientos de ancho de banda para video digital

Como ya se mencionó, la transmisión de video paquetizado se proyecta sobre una red de comunicaciones digital. Esto implica que la señal de video, que generalmente es producida por las cámaras de video en forma analógica, debe ser convertida al formato digital. Este proceso genera una determinada cantidad de datos digitales que es importante conocer para asignar el ancho de banda mínimo que se requiere para la transmisión eficiente del servicio a través de la red de comunicaciones.

El ancho de banda para video digital depende básicamente del espacio de color utilizado, la resolución temporal y la resolución espacial.

El espacio de color define la cantidad de unidades de información necesarios para representar a un elemento de imagen, expresado en bits/pixel. Como el estándar H.261 utiliza el espacio de colores YUV y considerando 8 bits por muestra, esto resulta en 12 bits/pixel en promedio, los cuales dependen del proceso de muestreo de las señales de luminancia y crominancia, en el formato 4:1:1.

La resolución espacial define el número de elementos que contiene una imagen o cuadro de video. De esta manera, para una imagen QCIF, el número de elementos por imagen resulta de $176 \times 144 = 25,344$ elementos. Otros formatos de imágenes de mayor resolución contienen un mayor número de elementos.

La resolución temporal indica la cantidad de elementos que deben transmitirse por unidad de tiempo, acorde a una cierta resolución espacial. Utilizando 30 cuadros/s como

resolución temporal y resolución espacial QCIF, la cantidad de elementos a transmitir se obtiene de $(25,344 \text{ elementos/cuadro}) \cdot (30 \text{ cuadros/s}) = 760,320 \text{ elementos/s}$.

Para obtener la tasa de transmisión o ancho de banda mínimo en bits/s, también se debe considerar al espacio de colores. Con el espacio de colores YUV, y para QCIF:

$$\text{tasa}_{\text{QCIF}} = (760,320 \text{ elementos/s}) \cdot (12 \text{ bits/elemento})$$

$$\text{tasa}_{\text{QCIF}} = 9.12 \text{ Mbps}$$

Cabe la observación de que el ancho de banda para video digital no comprimido depende únicamente de los tres parámetros analizados, y no depende del contenido de información de la secuencia, es decir, no importa si transmite videotelefonía, videoconferencia o secuencias de televisión comercial. También se observa que el proceso de digitalización produce inherentemente una tasa de transmisión constante.

En la tabla IV se muestra la tasa de bit requerida al transmitir video digital para diversos formatos de video, considerando 12 bits en promedio por elemento de imagen.

Tabla IV. Ancho de banda para diversos formatos de video.

Formato	Tasa de bit (ancho de banda) [Mbps]
QCIF	9.12
CIF	36.5
TV _{NTSC}	126
HDTV (propuesta americana)	330 (*)

(*) Podría utilizar el espacio de colores RGB con 24 bits en promedio por elemento, por lo que el ancho de banda crecería hasta 660 Mbps.

De la tabla III se puede concluir que la tasa de bit requerida para transmitir video digital es demasiado alta en comparación con la capacidad de los enlaces de acceso a los usuarios en los sistemas de telecomunicaciones actuales. Por ejemplo, para transmitir una imagen QCIF mediante una línea de acceso de 2 Mbps, se requiere un lapso de tiempo de 4.56 s, lo cual es intolerable para el servicio, ya que se desean 30 cuadros/s.

Evidentemente para formatos de más alta resolución a QCIF, la exigencia de mayor ancho de banda hace impráctico la disponibilidad de estos servicios. De aquí surge la necesidad de reducir la cantidad de datos generados por la digitalización de las señales de video analógicas, problema que se revisa en las siguientes secciones.

II.3 COMPRESIÓN DE VIDEO

Como se mostró en la sección anterior, la cantidad de datos generados por la digitalización de las señales de video resulta demasiado elevada, lo que provoca requerimientos imprácticos de almacenamiento, procesamiento y de comunicaciones, ya que unas cuantas fuentes de video o bien un solo servicio pueden saturar por completo la capacidad de los sistemas de comunicaciones actuales. Esto indica la necesidad imperante de reducir o comprimir esta cantidad de datos.

En esta sección se realiza un análisis descriptivo de las técnicas más utilizadas para tal propósito, enfocándose principalmente a los estándares de compresión de video.

II.3.1 Técnicas de compresión de imágenes

Para abordar el problema de comprimir video digital, es conveniente considerar a una sola imagen o cuadro, y posteriormente generalizar para una secuencia de imágenes transmitidas a una cierta velocidad.

La compresión de imágenes trata el problema de reducir la cantidad de datos requeridos para representar una imagen digital. La base fundamental del proceso de reducción es la remoción de datos redundantes [Gonzales y Woods, 1992].

Las técnicas de compresión se dividen en dos grandes categorías: preservación de la información (sin pérdidas) y con pérdidas de información.

La primera categoría es particularmente útil para imágenes de archivo, es decir, para almacenamiento de fotos legales o médicas, donde la compresión con pérdidas está prohibida por razones legales. Los métodos que caen dentro de esta categoría permiten a una imagen ser comprimida y descomprimida sin pérdidas de información. Sin embargo, la compresión de imágenes sin pérdidas rara vez resulta en una reducción de datos de más de 3:1 [Gonzales y Woods, 1992], lo que limita su utilización en la compresión de imágenes de video ya que se requieren mayores tasas de compresión.

Los métodos de la segunda categoría proveen más altos niveles de reducción de datos pero resultan en la reproducción menos perfecta de la imagen original. De esta forma, al contrario de la compresión sin pérdidas, la codificación con pérdidas está basada en el concepto de comprometer la exactitud de la imagen reconstruida a cambio de una mayor tasa de compresión. Si la distorsión resultante (la cual puede o no ser visualmente aparente)

llega a ser tolerable, el incremento en la compresión puede ser significativo. Estas tasas de compresión pueden llegar a ser de 30:1, 20:1, 10:1 para imágenes [Gonzales y Woods, 1992].

La compresión de imágenes con pérdidas es útil en aplicaciones tales como la difusión de televisión, videoconferencia y transmisión de facsímiles, en los cuales se aprovecha que una cierta cantidad de error es aceptable sin degradar en exceso la reproducción de las imágenes.

En la compresión de imágenes digitales tres tipos de redundancias básicas en los datos pueden ser identificadas y explotadas [Gonzales y Woods, 1992]:

- Redundancia de código
- Redundancia entre pixeles
- Redundancia psicovisual

La compresión de datos se lleva a cabo cuando una o más de esas redundancias son reducidas o eliminadas. A continuación se describen en que consiste cada una de estas.

II.3.1.1 Redundancia de código

La conversión de la señal de video analógica a formato digital asigna el mismo número de bits a todos los niveles de cuantización definidos. Esto generalmente produce un inconveniente ya que en muchas imágenes, ciertas tonalidades se presentan con mucha mayor frecuencia que otras. Esto significa que las tonalidades más frecuentes son codificadas de una forma que usan más símbolos de código que los absolutamente

necesarios. La imagen resultante se dice que contiene redundancia de código. En general, la redundancia de código se presenta cuando los códigos asignados a un conjunto de eventos (en este caso los niveles de tonalidades) no han sido seleccionados para tomar ventaja de las probabilidades de los eventos.

Si se asignaran menos bits a las tonalidades más probables que a las menos probables, se permite la compresión de datos. Este proceso es comúnmente referido como codificación de longitud variable.

II.3.1.2 Redundancia entre elementos de imagen

El proceso de codificación de longitud variable no altera el nivel de correlación entre los píxeles dentro de las imágenes. En otras palabras, los códigos usados para representar los niveles de tonalidad de cada imagen no tienen nada que ver con la correlación entre píxeles. Tal correlación resulta de la estructura o relación geométrica entre los objetos en la imagen.

Debido a que el valor de cualquier píxel dado puede ser razonablemente predicho del valor de sus vecinos, la información transmitida por píxeles individuales es relativamente pequeña. Mucha de la contribución visual de un solo píxel a una imagen es redundante; podría haber sido supuesta basándose en los valores de sus vecinos.

Este tipo de redundancia también se conoce como redundancia espacial, redundancia geométrica o redundancia intracuadro.

II.3.1.3 Redundancia psicovisual

El ojo no responde con igual sensibilidad a toda la información visual. Cierta información simplemente tienen menos importancia relativa que otra información en el procesamiento visual normal. Esta información se dice que es redundante psicovisualmente. Puede ser eliminada sin deteriorar significativamente la calidad de la percepción de la imagen.

En general, un observador busca distinguir características tales como esquinas y regiones de textura y mentalmente los combina en agrupamientos reconocibles. Entonces el cerebro correlaciona esos agrupamientos con el conocimientos previo para completar el proceso de interpretación de la imagen.

La redundancia psicovisual es fundamentalmente diferente a las redundancias discutidas anteriormente. Al contrario de la codificación y de la redundancia entre píxeles, la redundancia psicovisual es asociada a la información real o cuantificable. Su eliminación es posible sólo porque la información en si misma no es esencial para el procesamiento visual normal.

Ya que la eliminación de datos redundantes psicovisualmente resulta en una pérdida de información cuantitativa, ésta es comúnmente referida como una cuantización, ya que significa la transformación de un amplio intervalo de valores de entrada a un número limitado de valores de salida. Como es una operación irreversible (cierta información visual es perdida), la cuantización resulta en una compresión de datos con pérdidas.

II.3.2 Modelos de compresión de imágenes

Anteriormente se describieron tres técnicas para reducir o comprimir la cantidad de datos requerida para representar a una imagen. Sin embargo, dichas técnicas típicamente son combinadas para formar sistemas de compresión de imágenes prácticas.

Como se observa en la figura 5, un sistema de compresión consiste de dos bloques estructurales distintos: un codificador y un decodificador. Una imagen de entrada $f(x,y)$ es utilizada como entrada al codificador, lo cual crea un conjunto de símbolos de los datos de entrada. Después de la transmisión sobre el canal, la representación codificada es enviada al decodificador, donde una imagen de salida reconstruida $g(x,y)$ es generada. En general, $g(x,y)$ pudiera o no ser una réplica exacta de $f(x,y)$.

Tanto el codificador como el decodificador consisten de dos funciones o subbloques relativamente independientes. El codificador está compuesto de un codificador fuente, el cual remueve las redundancias de entrada, y un codificador de canal, que incrementa la inmunidad al ruido de la salida del codificador fuente. El decodificador incluye un decodificador de canal seguido por un decodificador fuente.

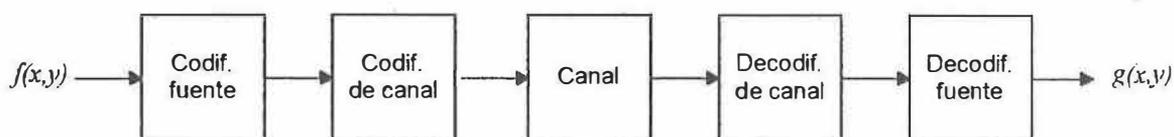


Figura 5. Etapas para un sistema de compresión de imágenes general

II.3.2.1 Codificador y decodificador fuente

El codificador fuente es responsable de reducir o eliminar cualquier redundancia de código, entre píxeles o psicovisual. La aplicación específica y los requerimientos de fidelidad asociados dictan el mejor procedimiento de codificación a usarse en una situación determinada. Normalmente el procedimiento puede ser modelado por una serie de tres operaciones independientes. Como muestra la figura 6, cada operación está diseñada para reducir una de las tres redundancias descritas anteriormente. También se muestra el correspondiente decodificador fuente.

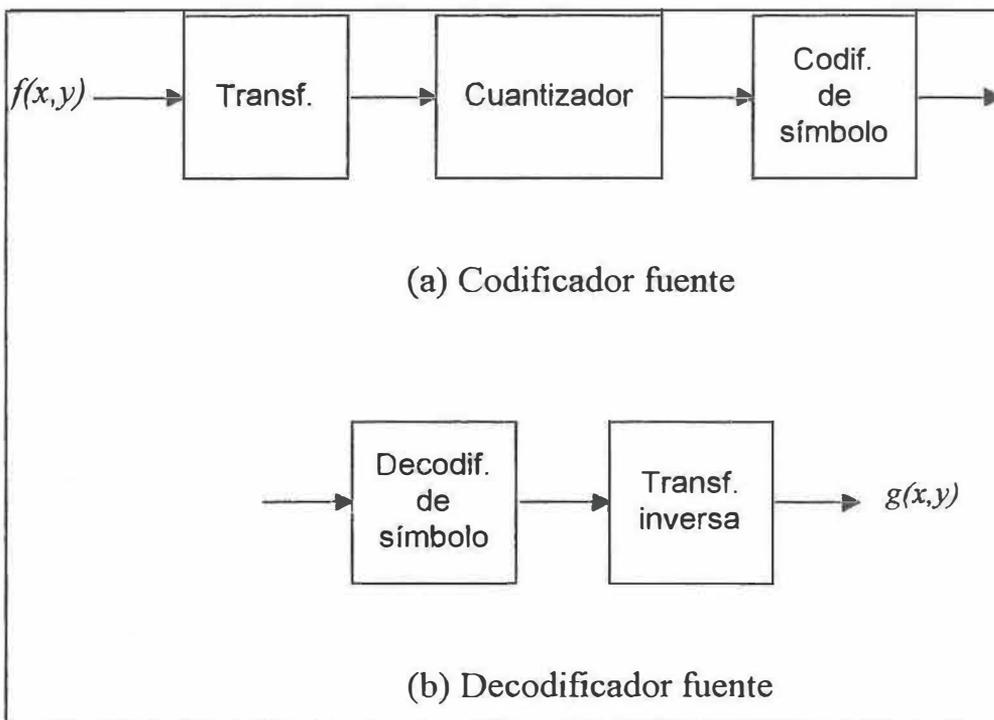


Figura 6. Procesos de (a) codificación y (b) decodificación fuente.

En la primera etapa del proceso de codificación fuente, los datos de entrada son transformados hacia un formato (usualmente no visual) diseñado para reducir las redundancias entre píxeles en la imagen de entrada. Esta operación generalmente es reversible y podría o no reducir directamente la cantidad de datos requeridos para representar a la imagen.

La segunda etapa o bloque cuantizador, reduce la precisión de salida de la transformación de acuerdo a algún criterio preestablecido de fidelidad. Esta etapa reduce la redundancia psicovisual de la imagen de entrada. Esta operación es irreversible y debe ser omitida cuando se desea compresión libre de error.

La tercera etapa del proceso de codificación fuente, el codificador de símbolo crea un código de longitud fija o variable para representar la salida del cuantizador y transforma la salida de acuerdo con el código. En muchos casos un código de longitud variable es usado para representar el conjunto de datos transformados y cuantizados. Asigna la palabra de código más corta a los valores de salida que ocurren más frecuentemente y así reduce la redundancia de código. Esta operación es reversible. Al final de la codificación de símbolo, la imagen de entrada ha sido procesada para remover cada una de las tres redundancias descritas anteriormente.

El decodificador fuente contiene sólo dos componentes: un decodificador de símbolo y un transformador inverso. Esos bloques realizan, en orden inverso, las operaciones de los bloques de codificación fuente, codificación de símbolo y transformador. Debido a que la

cuantización resulta en pérdidas de información irreversibles, un bloque cuantizador inverso no se incluye en el modelo decodificador fuente general.

II.3.2.2 Codificador y decodificador de canal

El codificador y decodificador de canal juegan un papel importante en el proceso completo de codificación-decodificación cuando el canal es ruidoso o propenso a errores. Son diseñados para reducir el impacto del ruido de canal al insertar una forma controlada de redundancia dentro de los datos entregados por el codificador fuente. Como la salida del codificador fuente contiene muy poca redundancia, podría ser altamente sensible al ruido de transmisión sin la adición de esta “redundancia controlada”.

El presente trabajo no aborda los problemas asociados al ruido en la transmisión, ni las técnicas de recuperación de errores más apropiadas para reducir sus efectos, así que una descripción mayor de las características del codificador y decodificador de canal se encuentran fuera del alcance de este trabajo.

Hasta este punto se abordó el tema de la compresión de imágenes fijas. Sin embargo, la codificación de video involucra algunos aspectos que no son aplicables en la codificación de imágenes fijas. La compresión de video se puede pensar como la generalización de los conceptos y técnicas desarrollados para la compresión de imágenes. Un proceso que resulta de la codificación de las secuencias de video es la aparición de las redundancias entre imágenes consecutivas o intercuadros, que modifica significativamente el esquema propuesto para la compresión de imágenes. La manera de definir como se realice la

reducción de redundancias intercuadro dependerá de la aplicación a la que se destine el sistema de compresión de video. De igual forma, es la manera exacta de implementar las técnicas de reducción de redundancias intracuarto. Esto conduce a la necesidad de revisar las técnicas empleadas por los estándares de compresión de video actuales y definir cual de estos es más apropiado para la parte experimental de este trabajo.

II.3.3 Estándares para compresión de video

A continuación se realiza una descripción de los estándares de compresión de video actuales, en función de sus aplicaciones y con la finalidad de determinar cual resulta más apropiado para el servicio de videoconferencia.

Los estándares formalmente adoptados para la compresión y descompresión de imágenes secuenciales son el H.261 (también referido como px64) de la ITU-T, y dos estándares adicionales llamados MPEG 1 y MPEG 2 del Grupo de Expertos de Imágenes en Movimiento de la ITU-T e ISO.

El estandar H.261 está proyectado para aplicaciones de videoteleconferencia, con retardos de transmisión de menos de 150 ms (los retardos que exceden 150 ms no ofrecen al observador la "sensación" de retroalimentación directa). En contraste, el estandar MPEG 1 se proyecta para aplicaciones con "calidad de entretenimiento" para almacenamiento y recuperación de imágenes comprimidas en medios digitales tales como una memoria de sólo lectura en disco compacto (CD-ROM). El estandar MPEG 1 está especificado para permitir mayor calidad de codificación que H.261. Sin embargo, no especifica un procedimiento de

codificación particular, meramente define un estándar para un flujo de bits codificados y su correspondiente decodificador. El estándar MPEG 2 soporta velocidades de transferencia de video entre 5 y 10 Mbit/s, un intervalo que es conveniente para distribución de TV por cable [Liou, 1991].

En lo particular, en este trabajo se decidió seguir la base planteada por el estándar H.261 debido a su mayor énfasis a la transmisión de señales de video en tiempo real y la característica de bidireccionalidad (interactiva) para la cual fue diseñada. La manera en la cual entrega la información este estándar debe estudiarse más ampliamente ya que se transmitirá dicha información por la red de comunicaciones bajo estudio. En el apéndice A se realiza un análisis descriptivo del esquema de codificación de video H.261.

Una vez seleccionado el estándar de codificación de video, resulta importante conocer el comportamiento temporal de la información generada por este sistema, es decir, la tasa de bit de video comprimido utilizando H.261, ya que esta información será la transmitida sobre la red de conmutación de paquetes. De igual importancia es conocer la calidad del servicio o la degradación causada por el sistema de codificación-compresión.

II.3.4 Características de la tasa de bit para video comprimido

Existen dos alternativas para la tasa de bit generada por los sistemas de compresión de video: a tasa de bit constante o tasa de bit variable. La selección por alguno de los codificadores depende de la capacidad de la red de comunicaciones de manejar tráfico de datos a velocidades variables [Ohta, 1994]. En la figura 7 se muestra el comportamiento

temporal de la tasa de bit al utilizar alguno de estos sistemas. También se observa el comportamiento ideal de la calidad del servicio a la tasa de bit correspondiente.

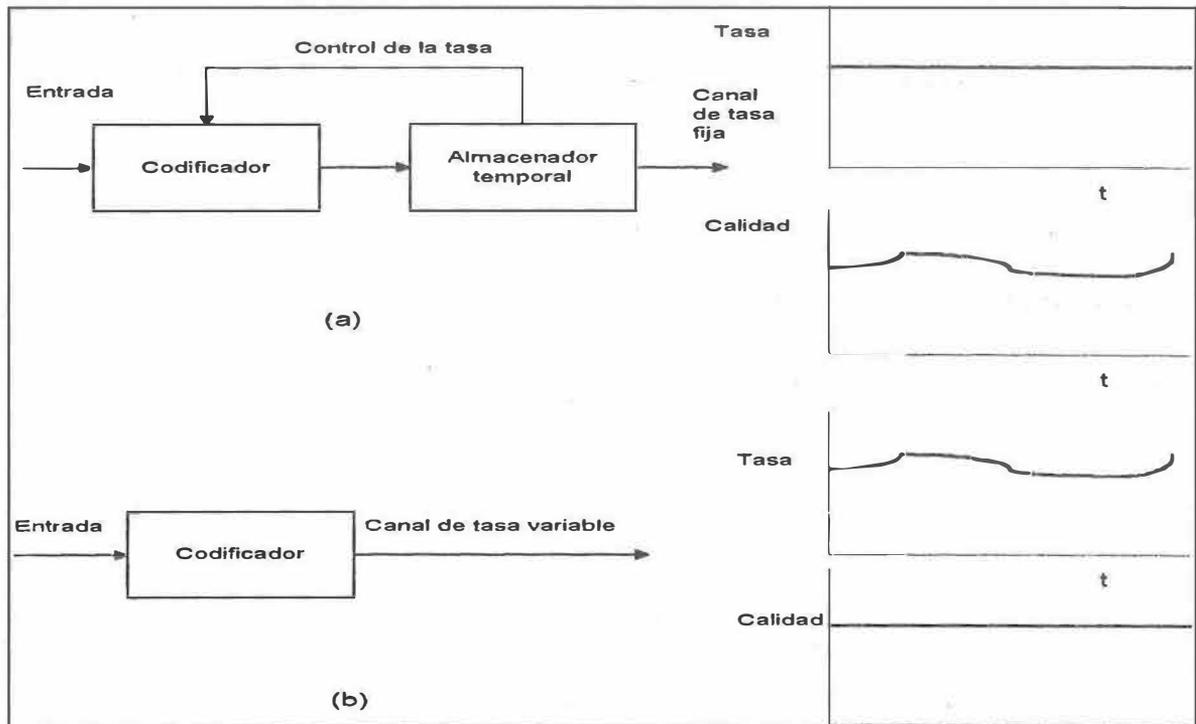


Figura 7. Tasa de bit y calidad generada por los sistemas de compresión de video: (a) tasa de bit fijo y (b) tasa de bit variable.

Existen algunos aspectos que es conveniente examinar en estas figuras. El primero se obtiene de la figura 7-(a), la cual muestra que la codificación a tasa de bit constante generalmente produce la degradación en la calidad del servicio, lo que teóricamente no sucede en la codificación a tasa de bit variable, ya que la calidad del servicio se mantiene constante.

Otro aspecto es la presencia de un sistema de regulación de la tasa de bit para el caso de la codificación a tasa de bit constante. Este sistema no aparece en la codificación a tasa de bit variable (figura 7-(b)). De hecho, las técnicas utilizadas por H.261 para comprimir las secuencias de video inherentemente producen códigos de longitud variable, los cuales al ser transmitidos generarían una tasa de bit variable. Sin embargo, H.261 fue diseñado para transmitir sobre una red basada en conmutación de circuitos, la cual no permite mucha flexibilidad para acomodar eficientemente el ancho de banda variable producido por el sistema de compresión. Por lo tanto, fue necesario incluir el sistema de regulación de la tasa de bit para permitir mejor utilización del ancho de banda [Karlsson y Vetterli, 1988; Lee y Wu, 1988].

En este capítulo se han descrito las técnicas más utilizadas actualmente en la codificación de video y las características que se aprovechan de las imágenes para reducir la cantidad de datos generados por el proceso de conversión a formato digital. De igual manera se han presentado la influencia de los sistemas de codificación en la tasa de bit para los servicios de video. Se determinó utilizar al esquema H.261 como referencia para el análisis de la transmisión de videoconferencia sobre la red de conmutación de paquetes y también en el estudio experimental para obtener datos reales en el proceso de codificación y compresión de secuencias de videoconferencia.

III DESARROLLO EXPERIMENTAL DE LA COMPRESIÓN DE VIDEO

En el presente capítulo se presentan las simulaciones realizadas para la compresión de video mediante el estándar H.261 y los resultados arrojados por tales experimentos. Una cuestión importante es determinar la degradación subjetiva inherente a los sistemas de codificación de video a tasa de bit constante. También se presenta un esquema modificado del sistema de compresión definido por H.261 con la finalidad de generar una tasa de bit variable y obtener así datos para el análisis de la transmisión de video a través de redes de conmutación de paquetes.

III.1 Esquema experimental de trabajo

En la figura 8 se observa la manera como se llevó a cabo la parte experimental del estudio de codificación y compresión de video. El esquema consistió de una secuencia de videoconferencia, un simulador del codificador H.261 el cual entrega como resultado (entre otros) el comportamiento de la tasa de bit medida en base a la cantidad de bits codificados por cada cuadro de la secuencia.

Como ya se ha mencionado, se utilizaron secuencias típicas de videoconferencia. En lo particular se seleccionó a la secuencia 'Salesman' sobre 'Miss America' como la secuencia de prueba debido a que la duración de la primera es mucho mayor, situación que permite mayor confiabilidad en las medidas estadísticas que serán útiles en capítulos posteriores del trabajo. El formato de video utilizado es QCIF tal y como es definido por el estándar H.261,

y la secuencia consta de 448 imágenes (la secuencia Miss America consta de 150 cuadros).

Estas secuencias se obtuvieron ya digitalizadas mediante Internet.

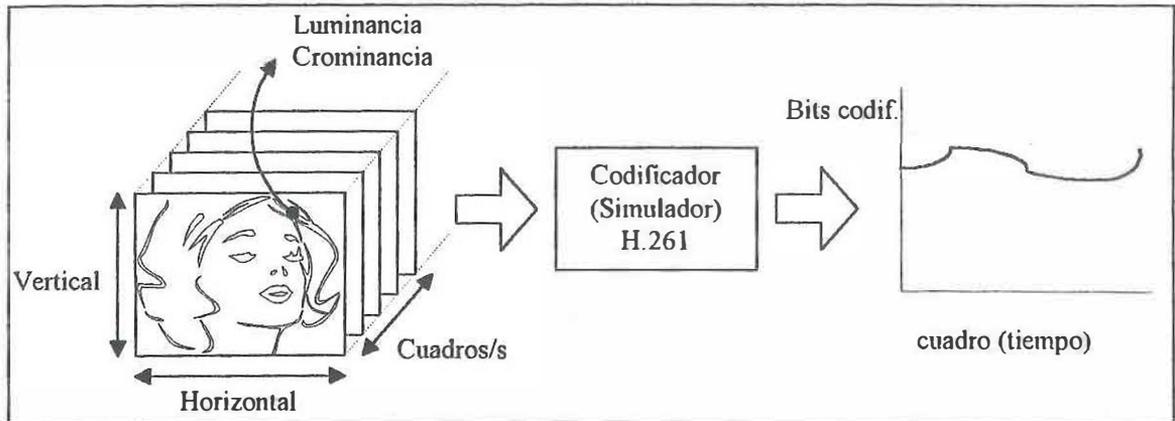


Figura 8. Proceso experimental para la compresión de video

Para el sistema de codificación H.261 se empleó el simulador p64 [U. Stanford, 1991] ya que resulta la manera más directa y económica de generar datos conforme a las especificaciones de la Recomendación H.261. El simulador se instaló y ejecutó en una estación de trabajo Sun Sparc 20. El proceso experimental se dividió en dos partes con la finalidad de cubrir las dos opciones para la transmisión de video paquetizado: a tasa de bit constante y a tasa de bit variable, y concluir sobre las ventajas y desventajas de ambos procedimientos de codificación.

III.2 Experimentos a tasa de bit constante

Para el caso de la transmisión a tasa de bit constante, es necesario asignar el algoritmo de regulación de la tasa de bit, el cual no es especificado por la Norma H.261. En el caso de el simulador empleado, este algoritmo se rige bajo el Modelo de Referencia 8 [Okubo, 1995], el cual se encuentra ya implementado dentro de la programática en el simulador.

Cada imagen está representada mediante 3 archivos ASCII diferentes, en los cuales se encuentra almacenada la información codificada en binario de las componentes de la señal de video. Uno de tales archivos define la luminancia Y y los otros dos representan la crominancia mediante U, V.

De la información generada por el simulador interesa particularmente la relacionada con las características de la tasa de bit y aquella que proporcione una medida de la posible distorsión a las imágenes causada por este codificador. Para el caso de las características de la tasa de bit, el simulador entrega una lista de valores relacionados con la cantidad de bits codificados por cada cuadro de video procesado. Sin embargo, para el caso de video comprimido a tasa de bit constante, el flujo de información a la salida del codificador es fijo. Por lo que el comportamiento de la tasa de bit es el mostrado en la figura 7-(a) del capítulo anterior.

Otro de los parámetros que deben de definirse al utilizar el simulador es la tasa de bit deseada a la salida del codificador. Con la finalidad de observar la influencia de la tasa de bit sobre la calidad subjetiva del servicio, se utilizaron una serie de valores dentro del intervalo permisible de 64kbps-2Mbps.

La figura 9 presenta algunas imágenes que resultan de los procesos de comprimir y descomprimir la secuencia "Salesman" a diferentes tasas de bit y a 30 cuadros por segundo. Las figuras 9-(a) y 9-(b) muestran el efecto sobre la calidad subjetiva cuando se codifica a bajas tasas de bit. La figura 9-(a) se codificó a 64 kbps y la figura 9-(b) a 128 kbps, tasas a las cuales está proyectado el servicio de videotelefonía. Es bastante notorio el efecto del sistema de compresión, ya que se observan los bloques sobre las imágenes, lo que provoca fuerte degradación en la calidad del servicio. Evidentemente, la degradación es más perceptible para la imagen codificada a menor tasa de bit.

Para los experimentos codificados a tasa de bit más alta, se observa que la degradación de la imagen disminuye notablemente. Por ejemplo, para la figura 9-(c) codificada a 380 kbps, tasa de bit a la cual se plantea transmitir videoconferencia de mediana calidad [Liou,1991], el efecto de los bloques no es perceptible por lo que la calidad de la imagen puede ser aceptable en este servicio. De igual manera se observa en la figura 9-(d) codificada a 1.5 Mbps, la cual no muestra degradación alguna. A este valor de la tasa de bit se considera adecuado para transmitir videoconferencia de alta calidad [Liou,1991].

Sin embargo, cabe la aclaración de que los experimentos se realizaron en base al formato QCIF el cual es considerado de baja resolución y fue originalmente planteado para aplicaciones de videotelefonía. Para la transmisión de videoconferencia es recomendable un formato de más alta resolución como CIF o inclusive uno de mayor resolución [Ohta,1995]. Los experimentos se llevaron a cabo en formato QCIF debido a la amplia disponibilidad de secuencias de prueba digitalizadas en este formato.

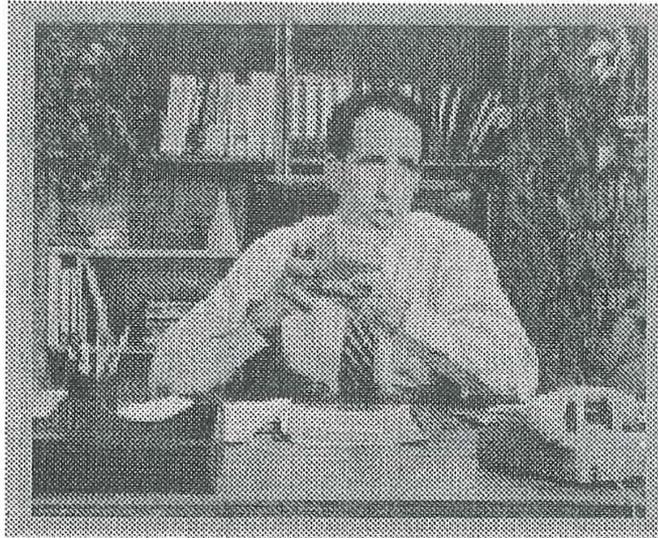


(a)



(b)

Figura 9. Efecto de la codificación a tasa de bit constante sobre la calidad de las imágenes.
(a) 64 kbps, (b) 128 kbps, (c) 380 kbps, (d) 1.5 Mbps.



(c)



(d)

Figura 9. Efecto de la codificación a tasa de bit constante sobre la calidad de las imágenes.
(a) 64 kbps, (b) 128 kbps, (c) 380 kbps, (d) 1.5 Mbps.

III.3 Experimentos a tasa de bit variable

Para obtener una tasa de bit variable, se empleó un algoritmo de compresión mediante transformación de coseno discreto combinado con codificación de longitud variable, siendo una modificación del estándar H.261 sin compensación de movimiento ni predicción intercadro con fines experimentales. Por lo tanto, el esquema de compresión solamente remueve redundancia espacial. La implementación del esquema de compresión propuesto fue mediante el simulador p64 para H.261. En el simulador se forzó el modo intracadro y el mecanismo de regulación de tasa de bit fue descartado.

A diferencia de la codificación a tasa de bit constante, la codificación a tasa de bit variable no utiliza un sistema de regulación que suavice las variaciones generadas por las técnicas de compresión empleadas por H.261, como se observa en la figura 7-(b) del capítulo anterior. En este caso es primordial obtener el comportamiento temporal de la tasa de bit generada al comprimir las secuencias de videoconferencia. Los datos más importantes obtenidos de la simulación que se requieren para analizar el comportamiento de la tasa de bit son la cantidad de bits codificados por cuadro de video. Esto implica que el sistema de compresión funciona en la base de cuadro a cuadro, es decir, el codificador toma una imagen, la codifica y comprime y posteriormente los datos resultantes son enviados al canal de transmisión. El proceso se repite de manera sucesiva. Considerando un determinado codificador y un formato de video particular, la cantidad de datos generados por cuadro codificado depende del movimiento de la escena y del contenido de información de cada imagen [Gonzales y Woods, 1992].

Para el esquema de compresión experimental a tasa de bit variable, fue necesario determinar un valor al cuantizador del sistema de compresión. Para ésto se hizo necesario codificar una imagen de la secuencia de video bajo prueba con diferentes valores del cuantizador y observar el resultado en la cantidad de datos codificados y la calidad subjetiva de la imagen resultante. En la figura 10 se observa el comportamiento de la cantidad de datos codificados para la primera imagen de la secuencia "salesman" cuando se varía el valor del cuantizador en el codificador de video.

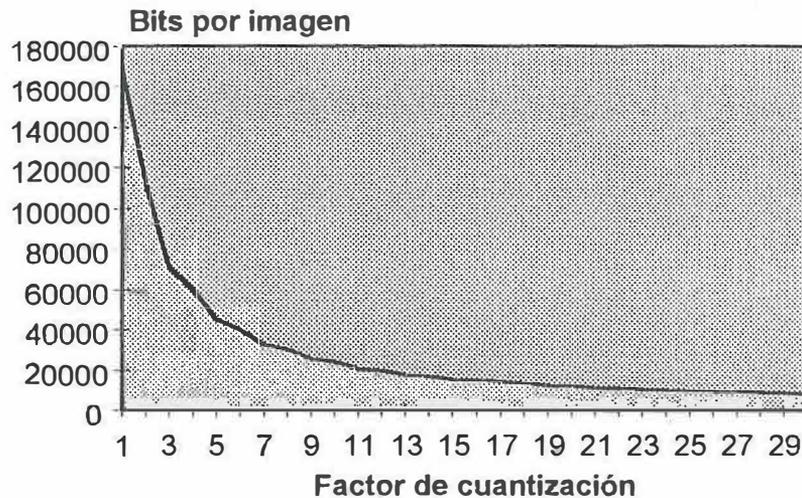


Figura 10. Reducción en la cantidad de datos comprimidos al modificar el valor de cuantización en un esquema de compresión de transformada de coseno discreto combinado con codificación de longitud variable.

De la figura 10 es notable la dramática reducción en la cantidad de datos codificados al incrementarse el valor del cuantizador. Si éste fuera el único parámetro que determinara el valor del cuantizador para los experimentos, sin lugar a dudas se escogería el valor más alto

(es decir 30) debido a que se generan menor cantidad de datos, lo que se reflejaría posteriormente en una menor tasa de bit a la salida del codificador. Valores más pequeños del cuantizador producen una cantidad mayor de datos codificados por imagen y, por lo tanto, generan tasa de bit más alta.

Sin embargo, en la figura 11 se muestran una serie de imágenes de la secuencia “salesman” obtenidas al codificarlas con algunos valores seleccionados del cuantizador, con la finalidad de observar la calidad subjetiva resultante.

De éstas se obtiene que al aumentar el valor del cuantizador se degrada la calidad de la imagen reconstruida. El efecto más notorio es el proceso de la transformación, ya que resulta evidente la aparición de las fronteras entre los bloques que se definen en éste esquema de compresión. Por lo tanto, no es conveniente utilizar un valor del cuantizador muy alto.

Considerando lo anterior, se eligió el valor del cuantizador de 10 para realizar los experimentos. En la figura 11-(b) se observa la calidad subjetiva cuando se comprime con este valor de cuantización. Aún cuando la reproducción no es perfecta, la degradación de la imagen no es demasiado evidente, por lo que se puede tolerar esta calidad de imagen. Por otro lado, se eligió porque no genera una cantidad excesiva de datos codificados, como se obtiene de la figura 10. Si bien, puede decirse que éste no es un procedimiento óptimo para la compresión de la secuencia de videoconferencia, si genera una degradación tolerable a la imagen tomada como referencia y además no produce una cantidad de bits demasiado alta comparado con los que producen otros valores de cuantización más bajos.



(a)

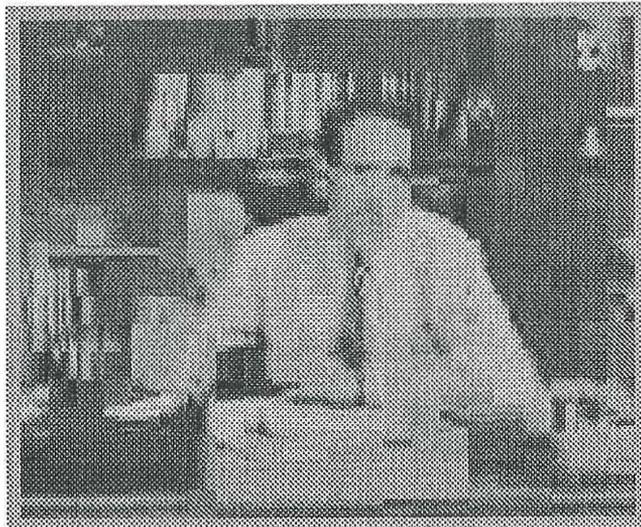


(b)

Figura 11. Imágenes decodificadas después de ser comprimidas con diversos valores de cuantización. (a) $q=1$, (b) $q=10$, (c) $q=20$, (d) $q=30$.



(c)



(d)

Figura 11. Imágenes decodificadas después de ser comprimidas con diversos valores de cuantización. (a) $q=1$, (b) $q=10$, (c) $q=20$, (d) $q=30$.

También de la figura 10 se observa que aumentar el factor de cuantización a valores mayores de 10, no resulta en una disminución significativa en la cantidad de bits codificados por cuadro, pero de acuerdo a las imágenes 11-(b) y 11-(d), la calidad subjetiva de las imágenes se degrada en forma notoria. Cabe hacer notar que la calidad de la secuencia decodificada se mantuvo para el resto de las imágenes que lo conforman, es decir, no hubo variaciones visibles en la calidad de las imágenes.

De acuerdo con este valor del cuantizador, se procedió a simular el proceso y posteriormente a obtener los datos buscados. En la figura 12, se presenta la tasa de bit $x(m)$ expresada en bits/cuadro, generada por la simulación anteriormente descrita.

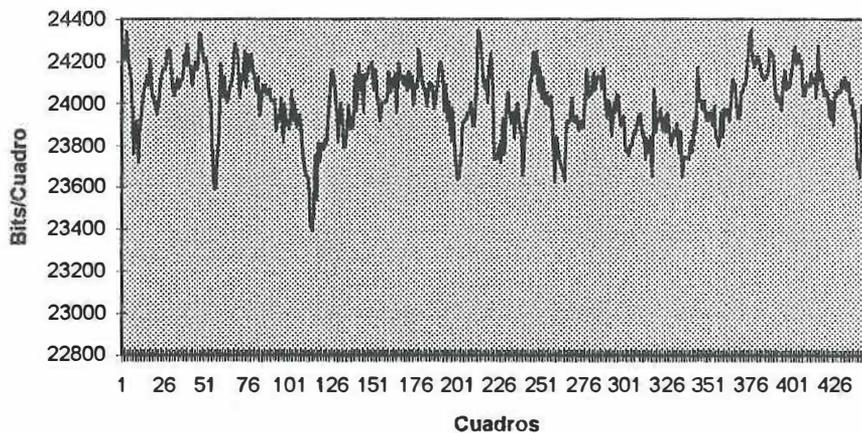


Figura 12. Tasa de bit generada por la secuencia "Salesman"

La variabilidad observable en la gráfica depende de la cantidad de movimiento presente en la secuencia. El valor promedio de la secuencia $\mu = E[x(m)]$, y la desviación estándar

$\sigma = E[x^2(m)] - \mu^2$ calculados resultaron $\mu = 24,000$ bits/cuadro y $\sigma = 164$ bits/cuadro. Cabe destacar que la relación $\sigma/\mu = 0.0068$ indica que realmente la variabilidad de los datos con respecto al promedio es bastante baja (es decir, ni siquiera uno por ciento en promedio). Esto se debe principalmente al escaso movimiento presente en la secuencia y a la codificación intracuadro (la mayor parte de las imágenes dentro de la secuencia la ocupa el fondo, el cual está siempre fijo, así que la codificación intracuadro tiende a generar la misma cantidad de datos por cuadro). Por esto, se obtiene una alta concentración de los datos alrededor del valor promedio, es decir, baja dispersión.

En resumen, en este experimento se permitió que la tasa de bit generada por el codificador variara de acuerdo al contenido de información de la secuencia de video, mientras que la calidad subjetiva se mantuviera constante. La determinación del valor del cuantizador tomó en cuenta el compromiso entre tasa de bit y calidad subjetiva. La calidad subjetiva se verificó visualmente al modificar el valor del cuantizador del codificador.

En este capítulo se presentaron los esquemas experimentales utilizados para obtener datos reales para el análisis de la transmisión de video paquetizado codificado a tasa de bit constante como variable. Estos también permiten obtener imágenes útiles para estimar subjetivamente la degradación del servicio a causa de la compresión con pérdidas. Se demostró que el esquema H.261 a bajas tasas de bit puede producir baja calidad del servicio. A medida que la tasa de bit se incrementa la degradación disminuye y por consiguiente la calidad aumenta. Todo resulta por efecto del sistema de regulación de la tasa de bit, que a su vez aparece por motivo de restricciones en las redes de comunicaciones. Sin embargo, una

de los inconvenientes de este sistema proviene de la variación en la calidad del servicio dependiendo de la actividad de la escena. Por otro lado, si se utiliza una tasa de bit lo suficientemente alta, la calidad del servicio puede mejorarse notablemente, como se puede observar en las imágenes 9-(c) y 9-(d).

Con las nuevas tendencias en las redes de comunicaciones, es posible transmitir información a tasas variables, por lo que resulta indispensable el estudio de los sistemas de compresión de video a tasa variable, que teóricamente ofrece la ventaja de calidad constante en el servicio de video. En la práctica es suficiente que el codificador genere una degradación máxima del servicio, i.e., mantenga una calidad de servicio mínima definida de acuerdo a las características de éste.

Otros experimentos [Ohta, 1994] han demostrado inclusive que la degradación del servicio puede resultar temporalmente inferior en los esquemas de codificación a tasa de bit constante con respecto a tasa de bit variable en momentos en los cuales el movimiento dentro de la secuencia de video es poco. Sin embargo, esto ocurre exclusivamente en ciertos instantes y no en toda la duración de la secuencia. Esto podría llevar a la conclusión de que la codificación a tasa de bit variable no conllevaría a la mejora de la calidad del servicio si tomáramos en cuenta exclusivamente los instantes en los cuales la calidad es superior cuando se utiliza la codificación a tasa de bit constante. Pero la codificación a tasa de bit variable ha demostrado que puede mantener un nivel mínimo de calidad durante todo el tiempo de transmisión, evitando las fluctuaciones en la calidad de las imágenes que son indeseables para el servicio.

Es necesario considerar otros aspectos para poder concluir en la superioridad de un esquema sobre otro, o en todo caso, obtener aplicaciones o condiciones en las cuales un esquema resulte más ventajoso que el otro. Entre estos aspectos a considerar están la vulnerabilidad de estos esquemas de codificación de video a los retardos de transmisión sobre la red de conmutación de paquetes, y la utilización eficiente del ancho de banda del sistema de comunicaciones definido para la transmisión de la información generada por cada uno de estos sistemas.

En el capítulo siguiente se aborda el primero de estos aspectos, considerando los dos esquemas de codificación de video planteados anteriormente. Por lo que respecta a este capítulo, se concluye que es preferible un esquema de codificación que mantenga una calidad mínima en el servicio. De acuerdo a los resultados obtenidos en esta parte experimental y apoyándose en los resultados de otros estudios anteriores, la codificación a tasa de bit variable presenta ventajas en esta característica con respecto a la codificación a tasa de bit constante.

IV ANÁLISIS DE RETARDOS PARA VIDEO PAQUETIZADO

Las aplicaciones multimedia integran una variedad de servicios, nominalmente, audio, video, imágenes, gráficas, texto y datos. Esos medios tienen diferentes requerimientos de ancho de banda, retardos y posibles pérdidas de información. El tráfico de datos es predominantemente de tipo a ráfagas y requiere servicios confiables por parte de la red, pero puede tolerar retardos. En contraste, el tráfico de audio y video es orientado a flujo de datos, es decir, un flujo continuo de información que debe ser entregado de manera oportuna. Puede tolerar algunas pérdidas de datos, pero, reiterando, requiere una entrega dentro de un retardo limitado. Los retardos de transmisión son irrelevantes para servicios de video unidireccionales, tales como televisión comercial. Para video interactivo como videoconferencia, en contraste, los retardos extremo a extremo son de importancia extrema, debido a que los retardos excesivos impiden el intercambio de información en tiempo real, situación última que da al usuario la sensación de interactividad.

Las aplicaciones interactivas tales como videoconferencia requieren que la suma de los retardos constantes y variables extremo a extremo del flujo de video sean menores de 100-200 ms [Dalgic, et. al., 1994]. Las aplicaciones interactivas no pueden comunicarse sobre redes que no garanticen retardos limitados. Los paquetes que experimenten retardos en exceso al retardo límite son considerados perdidos, ya que no serán utilizados en el proceso de decodificación en el extremo receptor. De esta manera, el reto para la red es ofrecer al servicio de videoconferencia retardos y pérdidas de paquetes limitados. Esto implica que es necesario estudiar las características de los retardos presentes en el sistema de

comunicaciones, ya que el proceso de video paquetizado involucra la transmisión de señales de video en tiempo real. También se requiere determinar cuales son los mecanismos inherentes en la codificación, transmisión sobre la red de paquetes y decodificación en el extremo receptor que provocan retardos en la señal de video para evitar, si es posible, que degraden a un nivel intolerable la calidad del servicio. Una vez identificados tales mecanismos, se debe investigar la posibilidad de obtener ciertos parámetros que los caractericen, con la finalidad de mantener bajo control los retardos producidos en cada uno de los procesos en donde éstos se originen.

IV.1 Identificación de retardos extremo a extremo para video paquetizado

Para el análisis siguiente los retardos extremo a extremo son aquellos producidos en la transmisión de video desde el codificador en el extremo transmisor, pasando por la red de conmutación de paquetes y hasta el decodificador en el extremo receptor. Durante toda esta trayectoria de transmisión existen dos tipos de retardos: fijos y variables. La premisa por un lado, es que la suma de estos retardos se mantenga bajo un cierto valor predeterminado (100-200 ms), el cual indique que la entrega de información se está llevando a cabo oportunamente en el extremo receptor, y por otro, controlar la variabilidad de dicho retardo. Si sucede lo contrario, la calidad del servicio es afectada considerablemente. De mayor importancia son los retardos variables, los cuales tienen mayor incidencia en la calidad del servicio si no son tratados adecuadamente en el receptor.

La figura 13 muestra a un panorama general de los elementos presentes en el proceso de transmisión de una señal de video paquetizado.

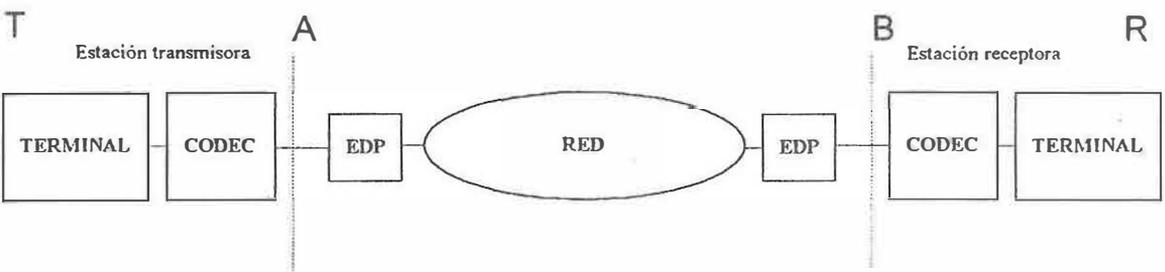


Figura 13. Elementos en la transmisión de video paquetizado en los cuales se producen retardos constantes y variables.

El estudio aquí presentado divide el análisis de retardos en dos partes: la primera en los aspectos de codificación y decodificación de la señal de video (CODEC). Esto se observa en la figura 13 marcado con los puntos T-A en el proceso de codificación (transmisión) y con B-R en la decodificación (recepción), los cuales se encuentran también marcados por líneas punteadas. La segunda parte considera la transmisión dentro de la red de conmutación de paquetes, marcado con los puntos intermedios A-B, en donde interesa conocer los retardos generados por el proceso de conversión de la información generada en la compresión de la señal de video a la unidad de transmisión de información definida por la red de paquetes (ensamblador-desensamblador de paquetes EDP), y el retardo variable producido por el almacenamiento temporal en el nodo de acceso a la red. Esta división del análisis se justifica debido a que la información de video sufre distintos retardos a medida que avanza

entre dispositivos desde su origen en la terminal transmisora hasta el desplegado en la terminal del receptor.

En la figura 14 se muestra el diagrama de tiempos extremo a extremo del flujo de video paquetizado.

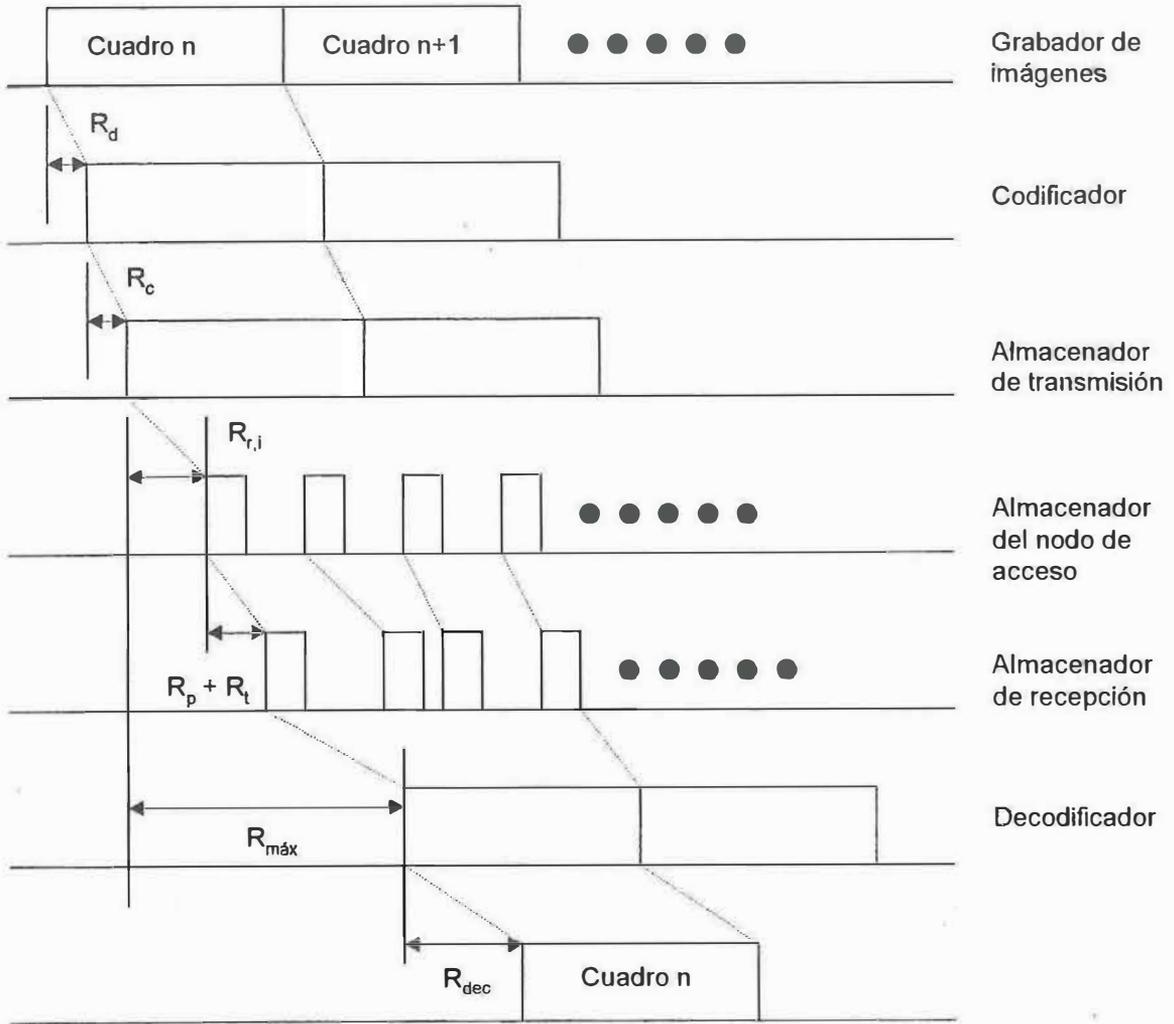


Figura 14. Diagrama de temporización del flujo de video

La figura 14 muestra que en el extremo transmisor (puntos T-A en la figura 13) los procesos que producen retardos se deben a la compresión de la señal de video, los cuales son básicamente la captura o digitalización de la imagen o cuadro (R_d), la codificación (R_c) y el almacenamiento temporal antes de enviarse a la red ($R_{r,i}$). Las características de estos retardos depende del esquema de compresión de video utilizado y de las características del canal (tasa constante o tasa variable). Posteriormente los datos generados por el codificador son paquetizados y transmitidos sobre la red de conmutación de paquetes (puntos A-B). Estos dos procesos producen retardos variables por efectos de paquetización (R_p) y almacenamiento temporal (R_t), mismos que se pueden compensar en el extremo receptor (puntos B-R) mediante el almacenador de recepción, que produce un retardo constante $R_{m\acute{a}x}$. Finalmente, los datos recibidos deben ser decodificados, lo cual también producen el retardo por decodificación (R_d).

A continuación se presenta el análisis para conocer los retardos constantes que se producen de extremo a extremo, y en base a ellos y al retardo máximo definido anteriormente para los servicio interactivos, obtener los intervalos de valores que pueden tomar los retardos variables para garantizar la interactividad en el servicio.

IV.1.1 Retardos de codificación y decodificación de la señal de video.

Este primer análisis considera los retardos generados en los procesos requeridos para comprimir la señal de video en el extremo transmisor y los correspondientes procesos

inversos que recuperan la información en el extremo receptor. Estos procesos están íntimamente ligados con el estándar de codificación y decodificación empleado y es conveniente analizarlos de manera conjunta. En este caso, el sistema de referencia es el ITU-T H.261, mismo que fue empleado en la parte experimental del trabajo del capítulo III. En tal capítulo se realizaron simulaciones para la transmisión a tasa de bit constante y también a tasa de bit variable, por lo tanto, a continuación se analizarán los retardos de procesamiento generado por ambos esquemas de codificación. También se discutirá las ventajas y desventajas de emplear alguno de estos sistemas de codificación en la transmisión sobre la red de paquetes.

IV.1.1.1 Codificación a tasa de bit constante

En este análisis el CODEC se considera que se encuentra apegado al esquema H.261.

Por lo tanto, las consideraciones acerca del funcionamiento del CODEC son las siguientes:

- a) El codificador solamente puede generar una tasa de bit constante, a la cual se le denomina v (misma que se encuentra en un intervalo seleccionable entre 64 kbps - 2 Mbps).
- b) El codificador procesa a la señal de video cuadro a cuadro como unidad básica de procesamiento.
- c) El decodificador puede esperar hasta que se reciba la información de un cuadro completo antes de que pueda empezar el proceso de decodificación.

Terminal y codificador.

La figura 15 muestra las fuentes de retardos para la codificación a tasa de bit constante. Este diagrama abarca la generación, codificación y almacenamiento temporal de la señal de video, marcado por los puntos T-A en la figura 13. Las fuentes posibles de retardos son: la cámara de video, el proceso de captura de imágenes, el codificador y el almacenador temporal de transmisión (del inglés buffering) [Dalgic, *et al.*, 1994].

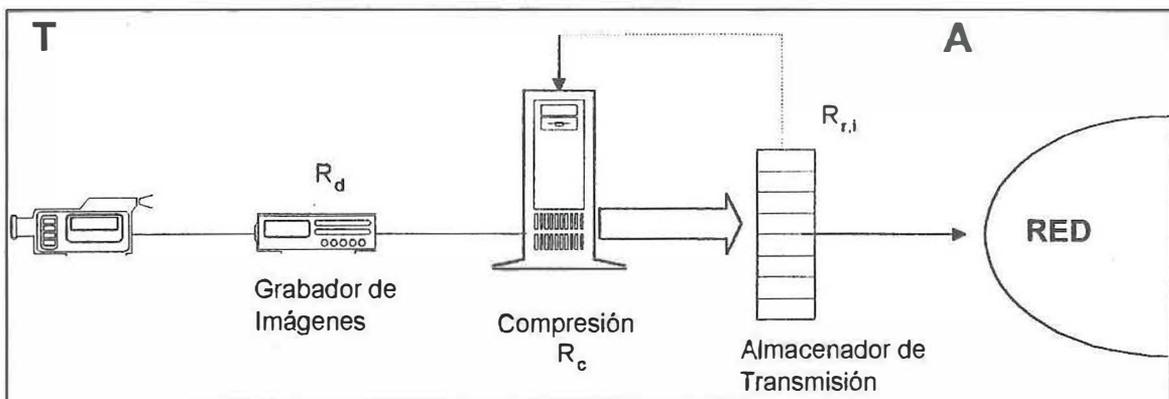


Figura 15. Fuentes posibles de retardos en el extremo transmisor de la señal de video.

- Cámara: entre la cámara y el capturador de imágenes (del inglés frame grabber) no existe retardo alguno, ya que el capturador digitaliza la señal analógica conforme ésta llega (procesamiento en tiempo real).
- Capturador de imágenes: se define R_d como el retardo en el capturador de imágenes. Aunque el capturador entrega la información digitalizada con retardo prácticamente cero

(proceso de conversión A/D bastante rápido), el codificador tiene que esperar hasta que se digitalice una imagen completa antes de iniciar el proceso de compresión. En este caso el tiempo de espera depende de la cantidad de cuadros de video que se transmitan. Como se ha mencionado en el desarrollo experimental, la tasa de cuadros transmitidos son 30 cuadros por segundo, lo que indica que el retardo producido en este proceso es $R_d = 1/30 = 33$ ms. Se supone que la cantidad de cuadros por segundo transmitidos no variará (no existen saltos de cuadros), así que este retardo es constante.

- Codificador: sea el retardo de codificación definido como R_c . Si el codificador opera en una base cuadro a cuadro, entonces $R_c \leq F$, donde F es el tiempo entre cuadros sucesivos.

Para el caso de 30 cuadros por segundo $F = 33$ ms, por lo cual $R_c = 33$ ms.

- Almacenador de transmisión: sea el tamaño del almacenador temporal B y el retardo de un cuadro en el almacenador temporal $R_{r,i}$. Entonces, el tiempo de permanencia de la información de cada cuadro variaría hasta un valor máximo de $R_{r,i} \leq B/v$. Esto significa que el retardo es directamente proporcional al tamaño del almacenador e inversamente proporcional a la tasa de bit utilizada para transmitir la información [Chen, *et al.*, 1993]. Este retardo es variable debido a que la cantidad de información codificada por cuadro es variable. Sin embargo, el estándar limita el número máximo de bits que pueden ser generados al comprimir un cuadro QCIF a 64 kbits y un cuadro CIF a 256 kbps. Tales límites definen el tamaño del almacenador temporal de transmisión y limitan el retardo de almacenamiento temporal.

Asumiendo la tasa de transmisión $v = p \times 64$ kbps, el retardo máximo será $1/p$ segundos para QCIF y $4/p$ segundos para CIF. Por lo tanto, el retardo generado en el sistema de regulación de la tasa de bit es variable, pero limitado a ciertos intervalos dependientes de la tasa de bit empleada. Cabe mencionar que este retardo puede resultar bastante grande, sobre todo en la transmisión a bajas tasas de bit (1 segundo para QCIF y 4 segundos para CIF considerando 64 kbps, 33 ms para QCIF y 133 ms para CIF al transmitir a 2 Mbps). Estos valores consideran casos extremos cuando se codifican el número máximo permisible de bits por cuadro. Sin embargo, un retardo típico para un codec H.261 es alrededor de 250 ms [Chen, *et al.*, 1993]. Este valor contrasta fuertemente con el retardo máximo permisible de 100-200 ms definido para los servicios interactivos y prácticamente no da margen para los retardos que se producen en la red de conmutación de paquetes. Esto significa que la transmisión de videoconferencia mediante el estándar H.261 está limitado sobre redes de comunicaciones que garanticen un retardo de transmisión mínimo al flujo constante de bits codificados.

Decodificación en el receptor

La figura 16 muestra los procesos en los cuales se presentan retardos en la terminal de recepción. Básicamente se generan en el almacenador de recepción y el proceso de decodificación.

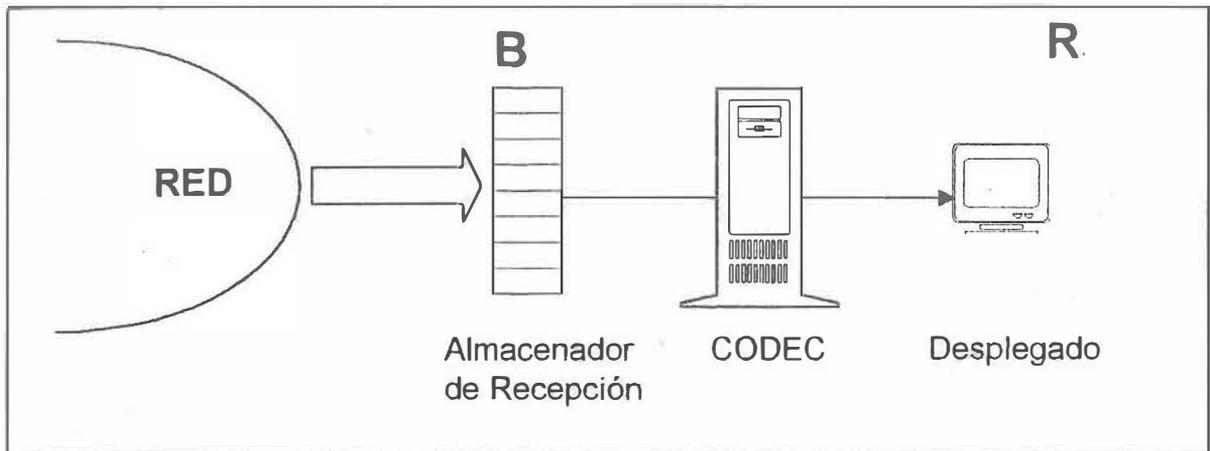


Figura 16. Fuentes de retardos en el extremo receptor de la señal de video.

En este extremo es necesario considerar que la información contendrá, además del retardo generado en el codificador, otras componentes de retardo variable provenientes de la red de conmutación de paquetes. La suma de los retardos variables producidos en la terminal transmisora más los retardos variables producidos en la red de conmutación de paquetes, pueden ser removidos al agregar un elemento de almacenamiento temporal en el receptor, por lo que en lugar de tener paquetes recibidos con retardos variables, el retardo se vuelve fijo e igual a un valor máximo [Karlsson y Vetterli, 1989; Lee y Wu, 1988].

- Almacenador de Recepción: la estación receptora almacena cada paquete recibido tal que el retardo desde la entrada al almacenador de transmisión hasta la salida del almacenador de recepción sea igual a un retardo máximo $R_{\text{máx}}$ ($R_{\text{máx}} > B/v$). Como ya se mencionó, la

finalidad de este elemento es introducir un retardo adicional constante para compensar las variaciones del retardo debido a la red y al sistema de regulación de la tasa de bit.

- Decodificador: sea el retardo de decodificación R_{dec} . Si el decodificador opera en base cuadro a cuadro, entonces como en el caso del codificador $R_{dec} \leq F$. Nuevamente para 30 cuadros por segundo, $R_{dec} = 33$ ms.

De esta forma, el retardo extremo a extremo se define como la suma parcial de los retardos producidos individualmente por todos los procesos involucrados en la transmisión [Dalgic, *et al.*, 1994]:

$$R_{total} = R_d + R_c + R_{m\acute{a}x} + R_{dec} \quad (1)$$

donde:

R_{total} : Es el retardo extremo a extremo

R_d : Retardo de captura

R_c : Retardo de codificación

$R_{m\acute{a}x}$: Retardo máximo de compensación

R_{dec} : Retardo de decodificación

Para los valores considerados, se obtiene que el retardo extremo a extremo en la transmisión de video a tasa de bit constante resulta de:

$$R_{\text{total}} = 99 \text{ ms} + R_{\text{máx}} \quad (2)$$

donde el valor para $R_{\text{máx}}$ debe considerar a los retardos variables producidos por el sistema de regulación de la tasa de bit y a la red de conmutación de paquetes. Si se toma el valor típico de 250 ms para el retardo en el almacenador de transmisión y se supone que la red no produce ningún retardo, el retardo total sería de $R_{\text{total}} = 300$ ms, valor que resulta bastante superior al esperado en las aplicaciones interactivas, además supone que la red de conmutación de paquetes no produce retardo alguno.

Este análisis para la codificación a tasa de bit constante, indica que el flujo de bits codificados por el estándar H.261 es extremadamente sensible a los retardos producidos en la transmisión. La razón fundamental radica en que su diseño está orientado a la transmisión sobre una red de comunicaciones basada en conmutación de circuitos, misma que no produce retardos considerables debido a su mecanismo de muticanalización y transporte. Por lo tanto, una red de conmutación de paquetes no ofrece la mejor alternativa para transmitir el flujo generado por H.261, debido a los retardos que inherentemente se producen en la misma.

A pesar de esto, es posible la transmisión de videoconferencia codificada por H.261 mediante una red de conmutación de paquetes si ésta asegura retardo mínimo en la transmisión mediante esquemas de prioridad de servicios, caso en el cual el retardo en el almacenador del nodo de acceso se pudiera anular ya que ese tipo de servicio tendría un

trato preferencial sobre el resto de la información que es multicanalizada en el nodo de acceso.

V.1.1.2 Codificación a tasa de bit variable

Para una red de conmutación de paquetes, un almacenador interno al codificador no es necesario debido a que la red puede absorber las fluctuaciones en la tasa de bit sin la necesidad de informar al codificador de decrementar la tasa como sucede para el procedimiento de regulación de la tasa de bit definido en el estándar H.261. Esto ha sido mostrado en la figura 7 del capítulo III. Además, la ausencia del almacenador en el codificador permite que retardos adicionales sean introducidos por la red de paquetes sin exceder el retardo total extremo a extremo [Lee y Wu, 1988].

A continuación se presenta un análisis similar al realizado en la sección anterior. De acuerdo a lo comentado anteriormente, el retardo por almacenamiento temporal no existe en el codificador a tasa de bit variable ($R_{r,i}$ en la figura 14). Solamente son aplicables los retardos por adquisición y codificación (figura 15) en el extremo transmisor. Por lo tanto, el retardo en la estación transmisora resulta de la suma de estos dos retardos, cuyo valor es de 66 ms constantes. Se observa que a diferencia de la codificación a tasa constante, el retardo disminuye notablemente y además es constante.

En el extremo receptor se aplica el mismo análisis que en la sección anterior. La única diferencia radica en que el retardo constante producido por el almacenador de recepción sirve para remover el retardo variable generado únicamente por la red de conmutación de

paquetes. Así que en el receptor el retardo producido es el de almacenamiento temporal y el de decodificación.

Si no hay pérdidas de paquetes en la red y suponiendo que no existen problemas de sincronización entre las terminales transmisora y receptora, entonces el retardo variable de los paquetes puede ser completamente eliminado al aplicar un tamaño de almacenador temporal suficiente (mayor o igual al retardo variable máximo de los paquetes) en la terminal receptora. De esta forma, el retardo variable de paquetes se convierte en un retardo fijo grande. Por lo tanto, mientras el retardo total esté dentro de límites aceptables el retardo variable puede ser completamente absorbido al asignar el almacenamiento temporal suficiente.

En resumen, en el receptor el retardo producido es el de almacenamiento temporal y el de decodificación. El retardo total estaría dado por:

$$R_{\text{total}} = R_d + R_c + R_{\text{máx}} + R_{\text{dec}} \quad (3)$$

De nueva cuenta, para los valores considerados, se obtiene que el retardo extremo a extremo en la transmisión de video a tasa de bit constante resulta de

$$R_{\text{total}} = 99 \text{ ms} + R_{\text{máx}} \quad (4)$$

donde el valor de $R_{\text{máx}}$ a diferencia de la transmisión a tasa de bit constante depende únicamente del retardo variable generado en la red de conmutación de paquetes. Esta expresión da el intervalo en el cual se permiten que fluctúen los retardos dentro de la red de conmutación de paquetes, si se considera que el retardo máximo permisible por el servicio de videoconferencia es de 200 ms y este valor se asigna a R_{total} , entonces el valor de $R_{\text{máx}}$ estaría dada por:

$$R_{\text{máx}} \cong 100 \text{ ms}$$

Por lo tanto, ésto indica que la suma de los retardos variables producidos por la red de conmutación de paquetes no debe de exceder 100 ms para asegurar que la interactividad del servicio permanezca y que el retardo variable pueda ser compensado en el receptor.

Si el retardo variable fuera el único fenómeno que afectara la transmisión de los paquetes en la red, entonces el problema de la transmisión de video paquetizado sería resuelto completamente mediante un almacenamiento temporal adecuado en el receptor, bajo las condiciones encontradas en el anterior análisis. Sin embargo, en muchos casos prácticos, las pérdidas de paquetes son inevitables en las redes de conmutación de paquetes. En este estudio se considera que la pérdida de paquetes puede deberse a dos causas fundamentales: una es el sobreflujo en el almacenador temporal del nodo de acceso a la red por lo que la pérdida de paquetes se debe al descarte de los mismos cuando existen problemas de congestión en la red, y la otra proviene del caso en la cual un paquete no llega

al receptor dentro del tiempo límite especificado en el almacenador destino, entonces este paquete es considerado perdido y un paquete de relleno es insertado para compensar este paquete perdido [Lee y Wu, 1988]. Por lo tanto, la red debe garantizar por un lado un retardo variable de paquetes máximo para asegurar que todos los paquetes lleguen a tiempo para ser decodificados en el receptor y por otro una máxima probabilidad de pérdida de paquetes que evite la degradación del servicio.

Para asegurar estas condiciones de operación de la red, es necesario analizar los procesos involucrados dentro de la red de conmutación de paquetes.

IV.1.2 Retardo dentro de la red de conmutación de paquetes

Haciendo referencia a las figuras 13 y 14, el retardo en los puntos A-B es el retardo producido en la red de conmutación de paquetes denotado como R_t . Este resulta de la suma de dos componentes, el tiempo de formación del paquete R_p y el retardo dentro de la red R_r que básicamente depende del proceso de multicanalización, almacenamiento temporal y envío sobre el canal de alta velocidad. Ambos retardos se clasifican dentro de los retardos variables. Sobre estos retardos, y de acuerdo al análisis de retardos extremo a extremo realizado para el caso de transmisión a tasa de bit variable, se obtiene que la suma de ambos debe de cumplir que:

$$R_t + R_p \leq R_{\text{máx}} \quad (5)$$

es decir:

$$R_l + R_p \leq 100 \text{ ms} \quad (6)$$

con la finalidad de asegurar que los paquetes sean entregados a tiempo en el decodificador.

Los retardos dentro de la red requieren un análisis detallado ya que se deben tomar ciertas consideraciones con respecto a las características de la red en cuanto a tamaño de paquete, velocidad de transmisión, esquema de multicanalización y conmutación, tamaño de almacenador, etcétera. Tal análisis se ve fuertemente impactado por las características de llegada de los paquetes al nodo de acceso a la red, es decir, de la tasa de bit generada por los codificadores de video. Por lo tanto, para llegar a conclusiones sobre el retardo en la transmisión es necesario realizar un estudio para caracterizar y modelar la tasa de bit generada por los sistemas de codificación de video. Debido a la extensión de tales análisis, éstos se presentan en los capítulos siguientes.

En este capítulo se ha presentado un análisis general de los retardos generados en el proceso de transmisión de video paquetizado a tasa de bit constante y variable. El análisis cubre el proceso de generación de la señal de video, su conversión a formato digital, la codificación y compresión y almacenamiento temporal anterior al envío hacia el canal de transmisión. También incluye el proceso inverso en el extremo receptor.

Los resultados muestran que la codificación a tasa de bit constante produce un retardo de codificación que pudiese llegar a valores demasiado altos (hasta de 250 ms típicamente), situación que limita su transmisión sobre una red de conmutación de paquetes, ya que ésta le agregaría otras componentes de retardos variables, que afectarían considerablemente a la

interactividad del servicio. Por lo tanto, la transmisión de video a tasa de bit constante sobre una red de paquetes es más conveniente si la red le otorga un trato preferencial sobre otros servicios, con la finalidad de producir un retardo mínimo a los paquetes generados por el codificador a tasa de bit constante, menor tasa de pérdida de paquetes, lo que permitiría disminuir de esta manera una posible degradación en la calidad del servicio.

Por su parte, el codificador a tasa de bit variable permite introducir retardos adicionales, diferentes a los producidos en la codificación, a los paquetes generados por este sistema, sin degradar necesariamente la calidad del servicio y tampoco recibir un trato especial por parte de la red de paquetes. Para este caso se obtuvo que la suma de retardos fijos y variables que pueden producirse en la red de conmutación de paquetes no deben de exceder de 100 ms con la finalidad de mantener la interactividad del servicio.

V CARACTERIZACIÓN DE LA TASA DE BIT PARA VIDEO PAQUETIZADO

En el capítulo anterior se concluyó que el retardo variable de transmisión más importante y que pudiese tener un efecto significativo en la calidad del servicio de videoconferencia se produce durante el proceso de almacenamiento temporal en el nodo de acceso a la red de conmutación de paquetes. El almacenamiento temporal forma parte del proceso de almacenamiento y envío que se lleva a cabo al multicanalizar la información proveniente de diversos servicios que deben ser transmitidos mediante la red de conmutación de paquetes. Como el proceso involucra una cola de espera de los paquetes en el almacenador temporal, entonces es dependiente de las características de arribo de la información al nodo de acceso. Por lo tanto, resulta indispensable caracterizar estadísticamente y modelar a la información que llega al nodo de acceso, es decir, la tasa de bit generada por los codificadores de video. Este modelo para la tasa de bit será útil en el análisis de colas (retardos) en el almacenador temporal.

V. 1 Caracterización para la codificación de video a tasa de bit constante

La codificación de video a tasa de bit constante genera un proceso completamente determinístico en la tasa de bit. De nuevo, haciendo referencia a la figura 7-(a) del capítulo II, se observa que la tasa de bit se caracteriza únicamente con el valor al cual está siendo codificada la señal de video. Esto significa que la información que arriba al nodo de acceso será siempre constante y cuyo valor es seleccionable en el intervalo de 64 kbps- 2Mbps. Este es el único parámetro que caracteriza a la información de video codificada a tasa de bit constante.

V.2 Caracterización para la codificación a tasa de bit variable

De acuerdo a los datos presentados en el capítulo III para el codificador a tasa de bit variable, la tasa de bit se representa por el proceso estocástico del número de bits codificados por cuadro de video. Este proceso se muestra en la figura 12 del capítulo III, la cual se presenta nuevamente a continuación por cuestiones de comodidad en la comparación.

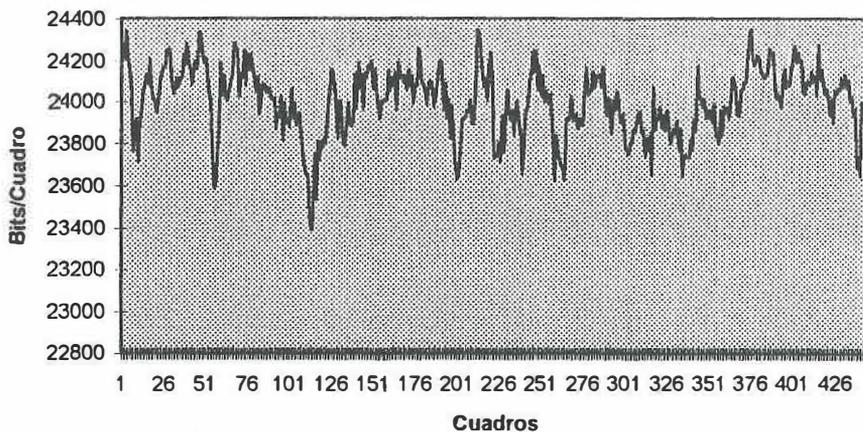


Figura 12. Tasa de bit generada por la secuencia "Salesman"

También en el capítulo III se presentan los valores estadísticos del número promedio de bits codificados por cuadro y la desviación estándar del mismo proceso. Existen medidas adicionales que permiten conocer más acerca de las características del proceso en cuestión, tales como la distribución probabilística y la función de autocovarianza [Maglaris, *et al.*, 1988].

Un histograma de la cantidad de bits codificados por cuadro se observa en la figura 17. Este indica la forma de campana de su función de densidad de probabilidad. Como se ha encontrado en otros estudios [Haskell, 1972; Maglaris, *et al.*, 1988; Ohta, 1994], la función de densidad de probabilidad del proceso de bits codificados por cuadro no resulta perfectamente simétrica alrededor de su promedio μ , a pesar de que se han utilizado otras secuencias de video y métodos de codificación diferentes al de esta simulación. En esta figura se observa también lo comentado con respecto a la relación σ/μ , ya que resulta notable la alta concentración de los datos alrededor del valor promedio.

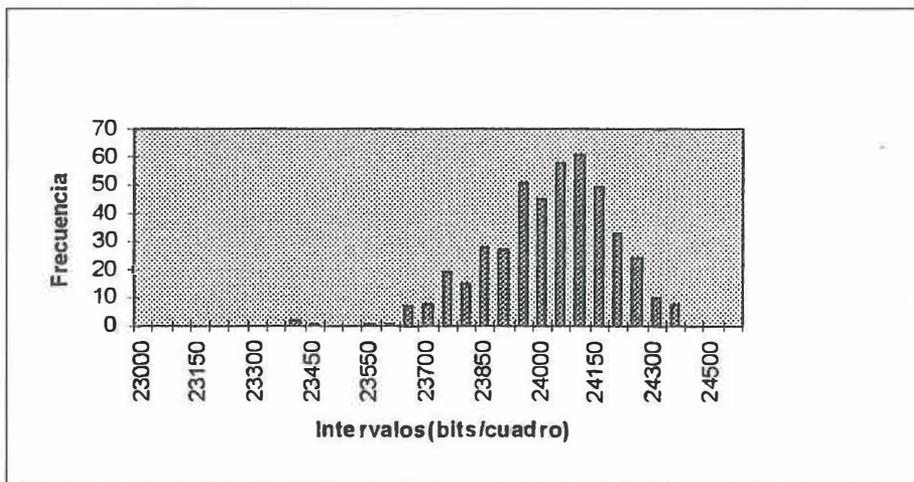


Figura 17. Histograma del proceso de bits codificados por cuadro .

En la figura 18 se presenta la autocovarianza $C(n) = E[(x(m) - \mu)(x(m+n) - \mu)]$ de los datos generados por la simulación, en función de la distancia entre cuadros n . Aquí cabe mencionar que $C(0)$ es la varianza de los datos calculada mediante σ^2 .

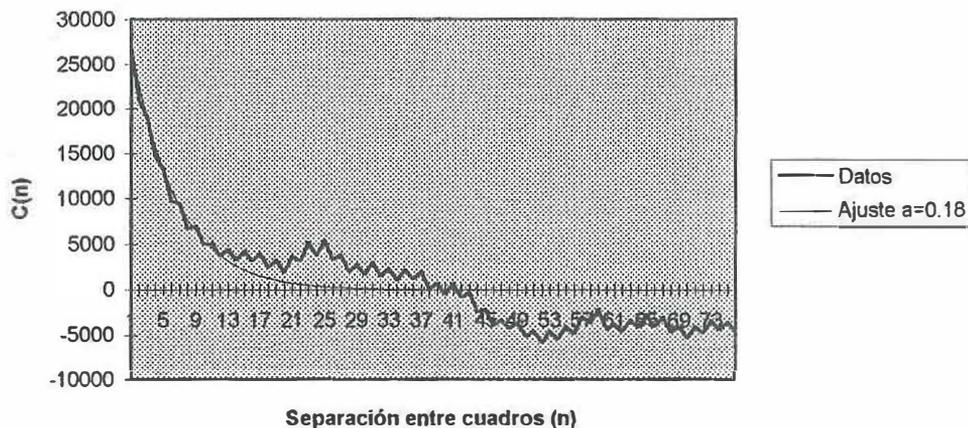


Figura 18. Función de autocovarianza para el proceso de bits codificados por cuadro.

Esta gráfica demuestra la forma exponencial de la autocovarianza de los datos. Mediante la línea delgada se incluye un ajuste exponencial de la forma $c(n) = \sigma^2 e^{-an}$ donde el valor de la constante a se calculó mediante aproximación sucesiva hasta que el ajuste exponencial se aproximara más a la autocovarianza calculada. El valor que más ajustó resultó $a = 0.18$ (mediante prueba y error). También se puede expresar el ajuste exponencial en el tiempo, considerando que $n = 30\tau$ (la frecuencia de cuadros por segundo), por lo que la autocovarianza resulta $c(\tau) = \sigma^2 e^{-5.4\tau}$. En este caso $a = 5.4$. El ajuste exponencial es casi exacto hasta que la diferencia entre cuadros llega aproximadamente a 13, en donde la autocovarianza medida es un poco mayor. También se observa un aumento considerable en la autocovarianza medida para n entre 20 y 30,

con un pico bastante notable para $n=25$. Más adelante, la autocovariancia medida decrece hasta valores menores a cero. Los valores negativos de la autocovariancia se deben a que los cálculos numéricos se realizan con muy pocos valores de la tasa de cuadros por segundo (es decir, el factor de autocovariancia es de orden n muy elevado en comparación a la cantidad de datos de que consta la secuencia de video) y podría ser mejorado si la secuencia de video fuera de mayor duración. En cambio, la diferencia observable para el pico presente en $c(\tau)$, es posible que se deba a que los movimientos realizados por la persona en la secuencia tienen un periodo de repetición natural de entre 25 y 30 cuadros, lo que en tiempo significa aproximadamente un segundo (movimiento de brazos y cabeza aproximadamente periódicos) [Maglaris, *et al.*, 1988 ; Ohta, 1994].

La autocovariancia exponencial también se ha encontrado en otros estudios para secuencias de videoconferencia y videotelefonía típicos sin cambios de escena aún con diferentes esquemas de codificación, como resultado de la alta correlación existente en el proceso de bits codificados por cuadro para las secuencias de videoconferencia o videotelefonía [Haskell, 1972 ; Maglaris, *et al.*, 1988; Verbiest y Pinno, 1988; Ohta, 1994].

En conclusión, se han calculado diferentes valores estadísticos de interés y, como se esperaba para este tipo de secuencias de videoconferencia, los resultados concuerdan con otros estudios similares: la tasa de bit no sufre cambios bruscos, su función de densidad de probabilidad resultó en forma de campana sin simetría con respecto al valor promedio y la autocovarianza presenta una distribución exponencial.

Partiendo de esta caracterización del proceso estocástico de bits codificados por cuadro, se requiere modelar este proceso para utilizar el modelo resultante en el análisis del retardo en el almacenador temporal.

V.3 Modelaje para la codificación de videoconferencia a tasa de bit variable

Antes de presentar el modelo empleado para la codificación de video a tasa de bit variable, se procederá a indagar la aleatoriedad del proceso al comparar los resultados obtenidos en la sección anterior, con la caracterización de una secuencia pseudoaleatoria.

V.3.1 Caracterización de una secuencia pseudoaleatoria

Un experimento adicional se planteó con la finalidad de simular una tasa de bit variable. Este consistió en generar una secuencia pseudoaleatoria con media μ_s y desviación estándar σ_s de igual valor a las obtenidas para los experimentos a tasa variable del capítulo III. Se desea determinar si tal secuencia posee las mismas características estadísticas de la tasa de bit variable generada por un codificador a tasa de bit variable. La secuencia pseudoaleatoria fue obtenida mediante el generador de números aleatorios del lenguaje de programación C++, con media $\mu_s=24000$ y $\sigma_s=164$ y generando 449 números pseudoaleatorios, las mismas cantidades de la secuencia de video "Salesman" codificada en modo intracuadro en el capítulo III. Cada número pseudoaleatorio representa en este caso la cantidad de bits codificados por cuadro.

La figura 19 muestra la secuencia pseudoaleatoria obtenida. El valor promedio resultó $\mu_s=24000$ y $\sigma_s=159$. Una comparación subjetiva entre la figura 12 del capítulo III, que se incluye de nuevo en este capítulo para hacer más fácil la comparación, y la figura 19, muestra que existen diferencias significativas en la forma de las curvas. La secuencia pseudoaleatoria tiende a cambiar de valor (alternar) significativamente entre datos consecutivos. Por su parte, la tasa de bit de la secuencia de video tiende a permanecer en valores próximos en muestras consecutivas.

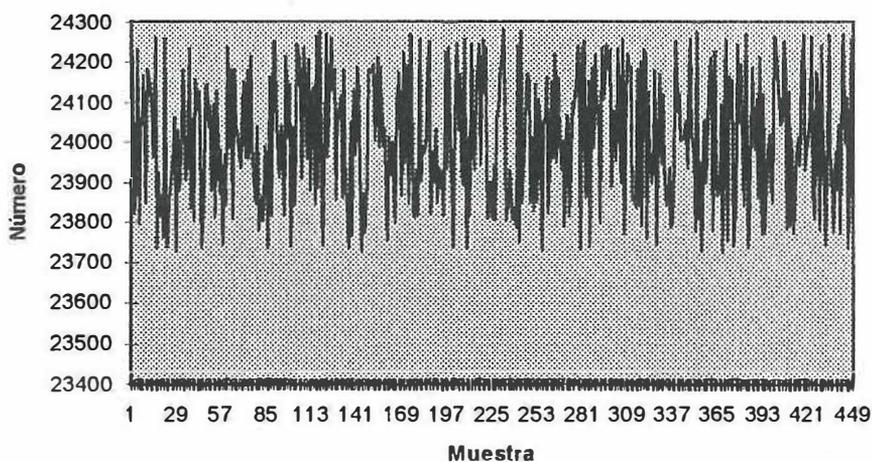


Figura 19. Secuencia pseudoaleatoria generada para analisis comparativo.

También se observan diferencias significativas entre los histogramas en las figuras 17 y 20. La tasa de bit posee una distribución en forma de campana, mientras que la secuencia pseudoaleatoria presenta una distribución uniforme.

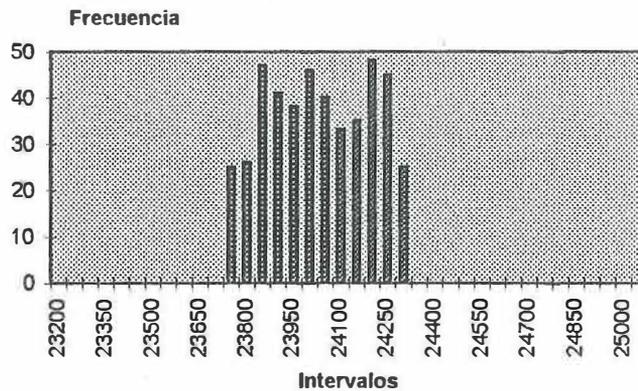


Figura 20. Histograma para la secuencia pseudoaleatoria.

La autocovarianza calculada a partir de la secuencia pseudoaleatoria se muestra en la figura 21. Esta no presenta una forma exponencial a diferencia de la figura 18. Se observa que la autocovarianza de la secuencia pseudoaleatoria se mantiene con valores cercanos a cero, es decir, la correlación no existe, a diferencia de la tasa de bit que posee cierta correlación por unos cuantos cuadros de video (al menos 13 de acuerdo a la figura 18).

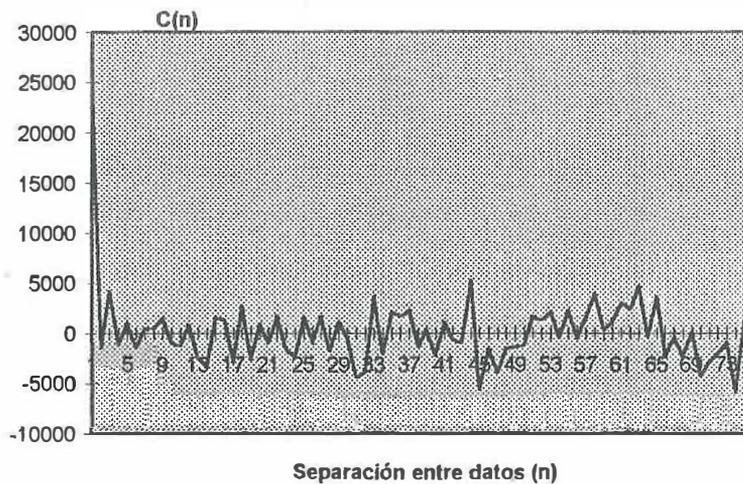


Figura 21. Autocovarianza de la secuencia pseudoaleatoria.

Para darle mayor validez estadística en la generación de la secuencia pseudoaleatoria, se incrementó el número de datos generados de 449 a 1,000, 2,000, 5,000 y 10,000. De esta manera los resultados estadísticos se obtienen de una población mayor y los resultados tienden a ser más confiables.

A continuación se presentan los resultados obtenidos para el caso de 10,000 datos. En la figura 22 se grafican los primeros 4,000 datos obtenidos para el caso del experimento con 10,000 números pseudoaleatorios.

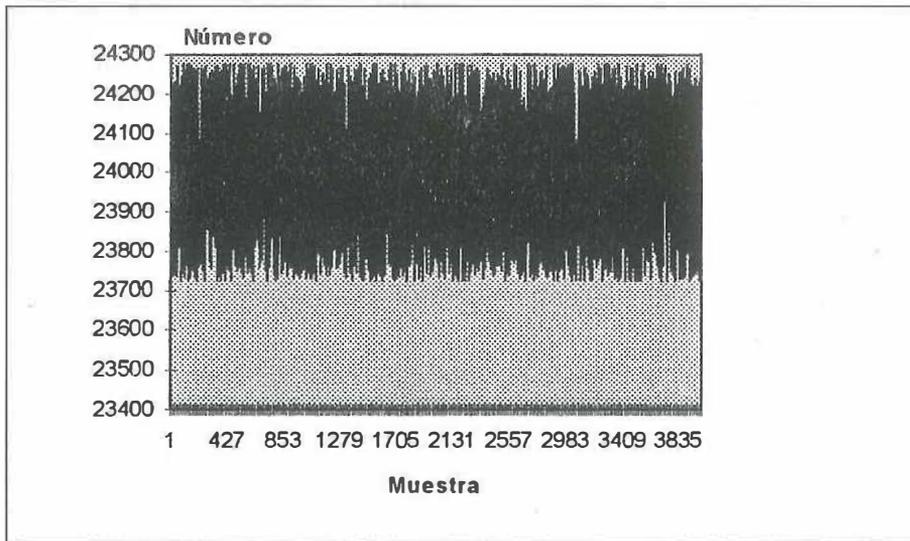


Figura 22. Secuencia pseudoaleatoria para 10,000 datos.

La figura 23 muestra el histograma para el caso de 10,000 números pseudoaleatorios. Se corrobora que la distribución de la secuencia generada tiende a ser más uniforme. Una distribución similar se obtuvo también para los otros casos de estudio con diferente cantidad de números pseudoaleatorios.

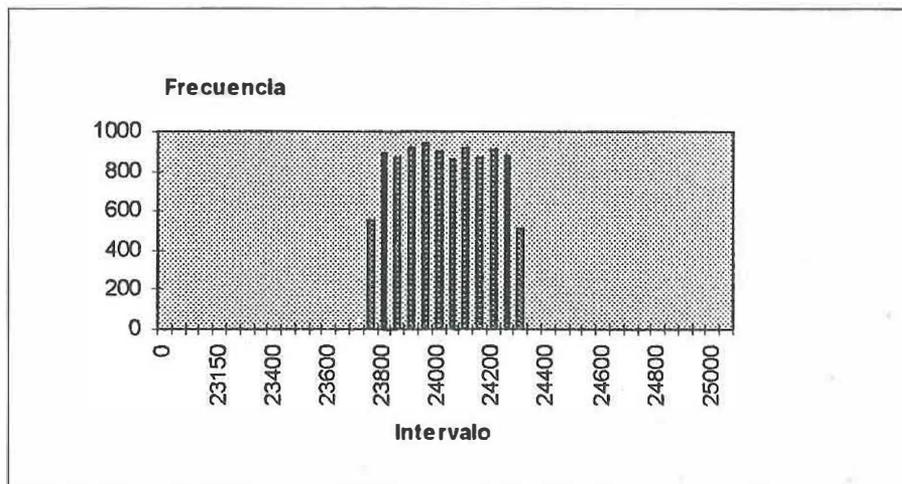


Figura 23. Histograma para el caso de 10,000 números pseudoaleatorios.

En la figura 24 se muestra la función de autocovarianza resultante para la secuencia pseudoaleatoria. El hecho de que la función de autocovarianza permanezca en cero es un indicativo de la ausencia de correlación entre los datos de la secuencia. Cabe mencionar que a medida que se aumentaba el tamaño de la secuencia de números aleatorios, la función de autocovarianza va tomando valores más próximos a cero (lo cual puede observarse al comparar las figuras 21 y 24) excepto para el primer valor de la función, el cual se mantiene prácticamente constante (la varianza de las secuencias no varía).

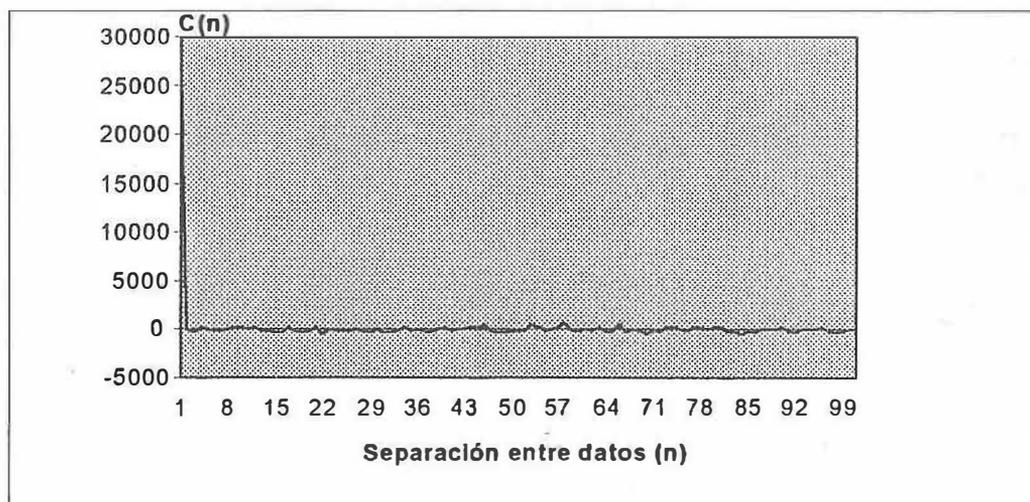


Figura 24. Función de covarianza para el caso de 10000 números pseudoaleatorios.

Estos experimentos llevan a concluir que, debido a las incongruencias existentes entre las características estadísticas de la secuencia de la tasa de bit experimental y la secuencia pseudoaleatoria, esta última (con el método empleado aquí para generarla) no es

una buena alternativa para modelar la tasa de bit de una fuente de video que transmite videoconferencia con las características ya descritas. La razón radica en que los procesos de compresión y codificación de video contemplado anteriormente y el algoritmo utilizado para la generación de números pseudoaleatorios son de naturaleza distinta y producen que los datos arrojados por cada uno de ellos posean comportamientos estadísticos completamente diferentes, siendo el más significativo la presencia de correlación para el primer proceso y su ausencia en el segundo.

V.3.2 Modelo para la tasa de bit variable

Para el análisis de multicanalización estadística de videoconferencia a tasa de bit variable el modelo presentado a continuación fue aplicado siguiendo los criterios establecidos por Maglaris [Maglaris, *et al.*, 1988]. Este es un proceso de Markov de estado discreto, tiempo continuo, donde la tasa de bit es cuantizada en niveles finitos discretos. Las transiciones entre niveles se suponen que ocurren con tasas de transición exponencial. La tasa de bit se aproxima con un proceso de tiempo continuo $x(t)$ con saltos discretos en tiempos de Poisson aleatorios. Este modelo puede ser pensado en un simil como una reserva ilimitada que está colectando agua de N fuentes de tasa variable, cada una con tasa $x(t)$. Sin embargo, el análisis se realiza sobre la tasa total agregada $x_N(t)$, en lugar de analizar la tasa de bit para cada fuente.

La tasa agregada resulta la suma de N procesos aleatorios independientes cada uno con media $E[x]$ y autocovarianza $C(\tau) \cong C(0)e^{-a\tau}$. Como son procesos independientes, la media y autocovarianza en estado estable de $x_N(t)$ resultan:

$$E[x_N] = N \times E[x] \quad \text{bits/cuadro} \quad (7)$$

$$C_N(\tau) = N \times C(0) \times e^{-\alpha\tau} \quad (\text{bits/cuadro})^2 \quad (8)$$

Los datos generados por la simulación no presentan discontinuidades abruptas debido a la ausencia de cambios de escena, como se mostró en la figura 12 del capítulo III. Como las variaciones en la tasa de bit son suaves, su suma no debe de exhibir cambios bruscos. En base a esto, se propone un modelo de Markov como el mostrado en la figura 25.

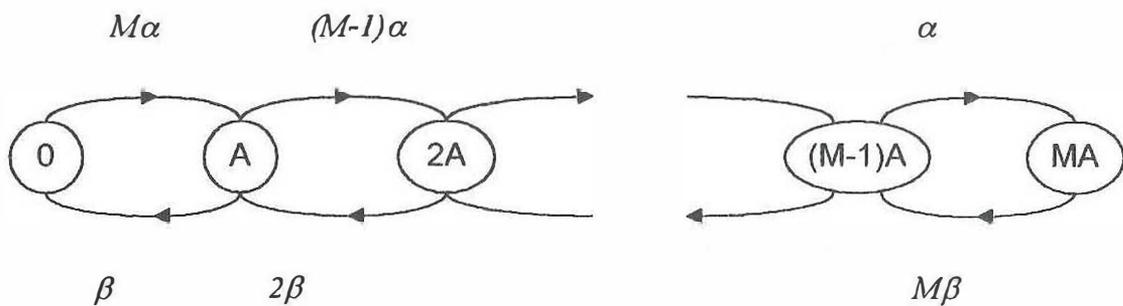


Figura 25. Diagrama de transición de estados para el modelo empleado

Este es un modelo de Markov del tipo nacimiento-muerte para describir la tasa de bit total, el cual sólo permite transiciones entre estados vecinos. En este modelo se espera que mientras la tasa de bit sea alta, su tendencia a seguir aumentando cada vez es menor y la tendencia a disminuir sea cada vez mayor. Si el nivel de la tasa de bit es baja, sucede lo contrario, ya que la tendencia de que disminuya la tasa de bit es cada vez menor mientras

que la tendencia a incrementarse es cada vez mayor. De esta manera, se espera que este modelo resulte en una distribución estacionaria de los estados en forma de campana, como la que se presenta en la figura 17.

En este modelo se asume niveles de cuantización uniformes de A bits/cuadro, y $M+1$ niveles posibles $(0, A, \dots, MA)$. Las tasas de transición exponencial r_{ij} del estado iA al estado jA están dadas como [Maglaris, *et al.*, 1988]:

$$r_{i,i+1} = (M - i)\alpha \quad i < M$$

$$r_{i,i-1} = i\beta \quad i > 0$$

$$r_{i,i} = 0$$

$$r_{i,j} = 0 \quad \text{abs}(i-j) > 1 \quad (9)$$

También se encuentra que el proceso $x_N(t)$ en estado estable tendrá una distribución binomial con media $E[x_N]$, varianza $C_N(0)$ y autocovarianza exponencial $C_N(\tau)$ dados por:

$$P\{x_N(t) = kA\} = \binom{M}{k} p^k (1-p)^{M-k}, \quad p = \alpha/(\alpha+\beta) \quad (10)$$

$$E[x_N] = Map \quad (11)$$

$$C_N(0) = MA^2p(1-p) \quad (12)$$

$$C_N(\tau) = C_N(0)e^{-(\alpha+\beta)\tau} \quad (13)$$

En este modelo se requieren conocer los parámetros M , A , α y β . Estos se obtienen al igualar los parámetros desconocidos con los datos estadísticos calculados. Dejando al número de niveles de cuantización M como un parámetro, y para un cierto número N de fuentes de video multicanalizadas, de acuerdo al modelo propuesto en [Maglaris, et. al., 1988] se tienen las siguientes expresiones:

$$\beta = a / (1 + (N \times E^2[x_N]) / (M \times C_N(0))) \quad (14)$$

$$\alpha = a - \beta \quad (15)$$

$$A = C_N(0) / E[x_N] + E[x_N] / M \quad (16)$$

Conociéndose de esta manera todos los parámetros del modelo, ya que a se obtiene del ajuste exponencial de la autocovariancia; $E[x(m)]$ y $C(0)$ fueron calculados con anterioridad. Solamente hace falta calcular los valores $E[x_N]$ y $C_N(0)$, los cuales pueden conocerse a partir de las expresiones (7) y (8).

Al sustituir los valores calculados para la secuencia simulada ($a = 5.4$, $E[x(m)] = \mu$ y $C(0) = \sigma^2$) en (7) y (8) se obtiene:

$$E[x_N] = 24000 \times N \quad \text{bits/cuadro} \quad (17)$$

$$C_N(\tau) = 26896 \times N \times e^{-5.4\tau} \quad (\text{bits/cuadro})^2 \quad (18)$$

Al sustituir (10) y (11) en (8), (9) y (10) resulta

$$\beta = 5.4 / (1 + 21416 \times N/M) \quad (19)$$

$$\alpha = 5.4 - \beta \quad (20)$$

$$A = 1.12 + 24000 \times N/M \quad \text{bits/cuadro} \quad (21)$$

pero resulta más conveniente expresar A en bits/s. Considerando que son transmitidos 30 cuadros/s

$$A = 33.6 + 720000 \times N/M \quad \text{bits/s} \quad (22)$$

En la figura 26 se observa la función de distribución de probabilidad para la tasa de bit total, resultante de sustituir los datos encontrados experimentalmente en (10), considerando $N=1$ (una fuente de video) y tomando como parámetro a M , es decir, variando el valor de cuantización A . El valor de k se encuentra dentro del intervalo $0, 1, 2, \dots, M$.

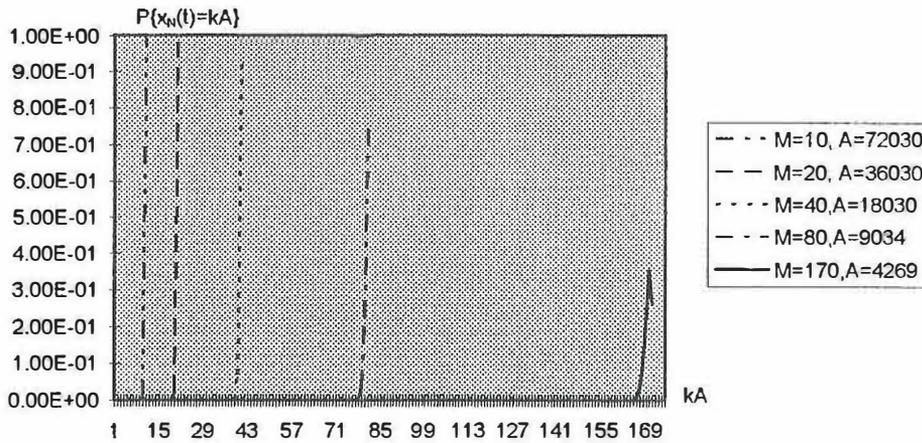


Figura 26. Probabilidades de la tasa de bit total del modelo, parametrizada por M .

Cabe mencionar que la figura 26 muestra 5 distribuciones de probabilidad distintas. Sin embargo, cada una de estas no presenta la forma de campana que se esperaba conforme a los datos experimentales y a la figura 17. Lo que se obtiene de la figura 26 es prácticamente una espiga por cada distribución de probabilidad. La mayor parte de las probabilidades calculadas resultaron cero o muy cercanas a cero, excepto para valores de kA cercanos al promedio, que es precisamente donde tiende a presentarse la "espiga". Esto se debe básicamente a dos cuestiones. La primera de éstas ya se ha expuesto anteriormente y resulta de la baja relación σ/μ de la tasa de bit experimental, es decir, los datos se encuentran bastante cercanos al valor promedio. La otra razón recae en el proceso de cuantización del modelo. Este proceso cuantiza desde el valor 0 bits/s hasta MA bits/s (se observa en el histograma de los datos experimentales, figura 17, una diferencia notoria: el proceso de la tasa de bit nunca toma el valor cero ni cercanos a cero), en intervalos regulares de A bits/s. A medida que M sea más grande, A tenderá a decrecer de acuerdo a la expresión (22). Conforme al intervalo de cuantización de la figura 18, que cuantiza en intervalos de 50 bits/cuadro, es decir, 1500 bits/s, y de acuerdo a (22) se requiere un valor de $M=491$ para conseguir que $A\cong 1500$ bits/s. Sin embargo, en (10) se requiere calcular el factorial de M , y el número que debe resultar supera por mucho la capacidad de cálculo de la aplicación de programación utilizada (Mathcad 5.0 realiza cálculos con exponentes hasta de 10^{307} , número que es superado cuando M apenas rebasa el valor de 170. Esto significa que solamente se pueden manejar hasta 170 niveles de cuantización para el análisis). En la figura 26 se observa para $M=170$, que la función de probabilidad apenas empieza a tomar una cierta forma de campana, pero debido a que

el valor de cuantización $A=4269$ bits/s es de la misma magnitud que la desviación estándar $\sigma=4920$ bits/s, no es posible obtener completamente la forma de campana del modelo para la tasa de bit total.

Por otra parte, la figura 27 muestra la autocovarianza del modelo para una sola fuente. La curva es idéntica al ajuste exponencial hecho a la autocovarianza experimental. Esto resulta de que el modelo inherentemente posee autocovarianza de tipo exponencial, así que existe congruencia para esta característica entre los datos experimentales y el modelo.

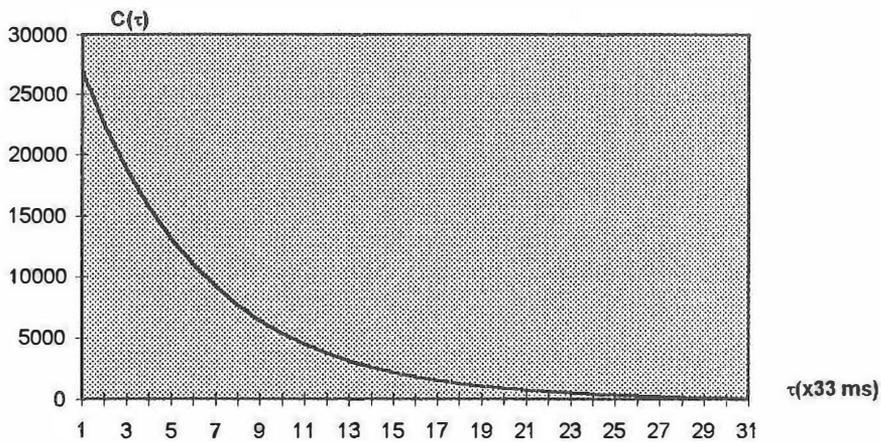


Figura 27. Autocovarianza de la tasa de bit total. $N=1$.

Como conclusión, en esta segunda sección, correspondiente al modelaje de la tasa de bit variable, se ha adaptado y aplicado un modelo para analizar la tasa agregada de N fuentes de video. Diversos parámetros estadísticos del modelo se han ajustado a los correspondientes valores experimentales, y como se mostró, el modelo funciona bien. Sin embargo, existe una limitación por la manera en la cual ha sido definido. Esta resulta en la

definición de los niveles de cuantización del modelo, ya que empieza en cero y, como pudo observarse, la tasa de bit experimental no toma valores cercanos a cero. Esto redundaba en que se debían definir un número grande de niveles de cuantización (M) de tal forma que los intervalos de cuantización A sean lo suficientemente finos como para permitir modelar con precisión la tasa de bit experimental. En lo particular, debido a que la relación σ/μ experimental resultó muy baja, M debería ser bastante grande. Sin embargo, debido a limitaciones de la programación utilizada para realizar los cálculos, no se pudo aproximar con suficiente exactitud la función de probabilidad del modelo. Por lo tanto, si se pretende mayor precisión en el modelaje, es necesario utilizar una herramienta computacional que posea un potencial mayor en los cálculos.

El análisis anterior ha mostrado que el modelo para la tasa de bit agregada resulta satisfactorio. Sin embargo, en la figura 26 no se pudo mostrar la forma de campana de la distribución de probabilidad del modelo. Sin embargo, para demostrar que dicha distribución sí posee la forma mencionada, se propuso una serie de experimentos que consisten en modificar arbitrariamente el valor de σ y mantener constante μ , de tal manera que se modifique la distribución de probabilidad de acuerdo a (22). La razón de elegir a σ es para que varíe la relación σ/μ , ya que se ha supuesto que la relación σ/μ obtenida para los datos experimentales es una de las causas de la aparición de “espigas” en la distribución de probabilidad del modelo.

Los valores propuestos para σ son $\sigma_1=10\sigma$, $\sigma_2=1/4\times\mu$, $\sigma_3=1/2\times\mu$. El valor de a no se modificó y M se mantuvo al valor de 20 para los tres casos. La figura 28 muestra las distribuciones de probabilidad obtenidas para estos experimentos. Los valores para σ y A

se expresan en bits/s. La forma de campana es evidente. Para todos los casos A resultó menor que σ , situación que no sucedió para los datos experimentales, y tales valores de A sí permiten definir con nitidez la forma natural de la distribución del modelo. Sin embargo, en estos experimentos, el programa de aplicación no tuvo problemas con el valor de M para realizar los cálculos de la distribución, ya que A no requería ser demasiado pequeño.

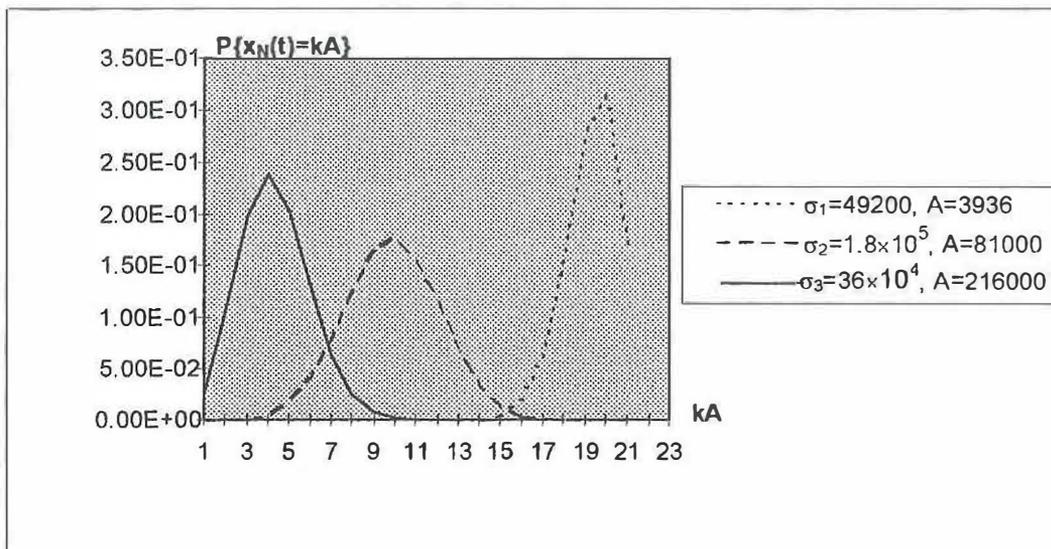


Figura 28. Distribuciones de probabilidad para diversos valores de σ .

Estos experimentos permiten ratificar la validez del modelo, ya que representa adecuadamente las medidas estadísticas obtenidas de los datos experimentales: distribución en forma de campana y autocovarianza exponencial. Ratifica también que los resultados presentados en la figura 26 se deben a limitaciones en la capacidad de cálculo de los programas de aplicación utilizados para el análisis y no a limitaciones propias del

modelo. Por lo tanto, se concluye que este modelo es adecuado para la tasa de bit generada por videoconferencia a tasa de bit variable, tal y como se presentó en el capítulo III lo cual justifica su uso en el análisis de retardos de almacenamiento temporal en la red de conmutación de paquetes que se realizará en el capítulo siguiente.

VI ANÁLISIS DE RETARDOS EN LA RED DE CONMUTACIÓN DE PAQUETES

Como ya se mencionó, en la transmisión a través de la red de conmutación de paquetes existen un par de procesos que producen retardos variables, los cuales resulta necesario analizar. Estos retardos provienen del proceso de la formación del paquete y del tiempo de espera en el almacenador temporal del nodo de acceso y dependen de las características de la tasa de bit generada por los codificadores de video. Por lo tanto, el análisis de cada uno de estos retardos se divide, de acuerdo a la tasa de bit generada por los codificadores de video, en dos partes: codificación a tasa de bit constante y codificación a tasa de bit variable. Sin embargo, se ha demostrado la conveniencia de que la transmisión de video paquetizado a tasa de bit constante se realice mediante esquemas de prioridad hacia este tipo de flujo de información. El modelo del nodo de acceso presentado en este capítulo considera los aspectos de almacenamiento y envío mediante un orden primero que entra-primero que sale, sin esquemas de prioridad de servicio, ya que se encuentra fuera de los alcances del trabajo. Por lo tanto, el análisis de almacenamiento y envío solamente se efectúa para el caso de tasa de bit variable. El análisis de paquetización se realiza para ambos esquemas de codificación.

VI.1 Retardo generado por la paquetización

La paquetización es el proceso de conversión o transformación de los datos generados por las fuentes de video hacia la unidad básica de transmisión (paquete) empleada por la red de conmutación de paquetes. Este proceso no es inmediato, sino que

requiere un cierto tiempo de procesamiento por parte del ensamblador-desensamblador de paquetes. Además, es necesario que esta interfaz espere la llegada de los datos para formar tales unidades o paquetes, produciéndose aparte del procesamiento un retardo de espera. Este retardo de espera generalmente es mucho mayor (dominante) que el retardo de procesamiento y para el presente trabajo es el único que se considerará.

Como el retardo de espera depende de la tasa de arribo de datos y de las características de la unidad básica de procesamiento de la red, es necesario definir ambas cosas para establecer la validez del siguiente estudio.

La red define como unidad básica de transmisión de información a un paquete corto y de longitud fija llamado celda. La celda consta de 53 octetos, 5 de los cuales forman el encabezado y los 48 restantes se toman como la carga útil. En estos 48 octetos se transmitirá la información neta de video.

Por lo que respecta a la tasa de arribo de los datos, como ya se mencionó, el estudio considera los dos esquemas : para video a tasa de bit constante y video a tasa de bit variable.

VI.1.1 Fuente de video con tasa de bit constante

En este análisis se considera la transmisión de video codificado mediante el estandar H.261 a través de una red de conmutación de paquetes y se supone que la tasa de bit es constante a una velocidad de $p \times 64$ kbps, donde p puede tomar un valor entero entre 1 y 30.

Aquí se supone que los $px64$ kbps son exclusivamente información de video. No existen datos provenientes de otros servicios (voz) intercalados con el video (multicanalizado). Tampoco se considera la existencia de bits de relleno.

En la figura 29 se muestra el proceso de formación de paquetes en el ensamblador-desensamblador de paquetes (EDP). Al EDP está arribando información a tasa de bit constante, con las características ya mencionadas. El EDP espera hasta que arriben mínimamente 48 octetos de información necesarios para formar una celda, produciéndose un retardo de espera R_p entre celdas consecutivas.

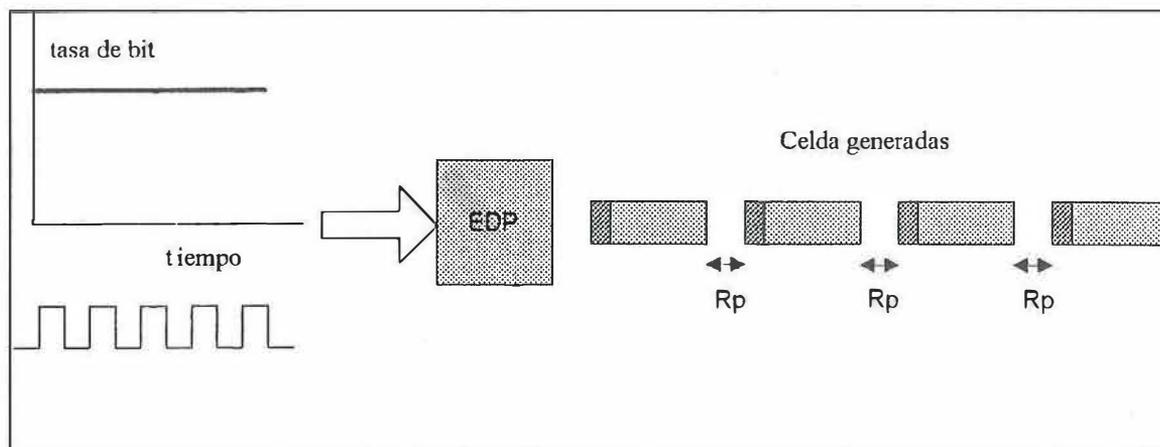


Figura 29. Retardo producido por el proceso de formación de la celda a tasa de bit constante

El retardo resulta del tiempo requerido para “llenar” la porción de la carga útil de la celda. Se define por R_p , y es directamente proporcional a la longitud de la carga útil de la celda en octetos, e inversamente proporcional a la velocidad de arribo de los datos de video. Se calcula como:

$$R_p = \frac{(48 \text{ octetos}) * (8 \text{ bits / octetos})}{p * 64 \text{ kbps}} \quad (23)$$

$$R_p = \frac{384}{p * 65536} \quad \left[\frac{\text{seg}}{\text{celda}} \right] \quad (24)$$

Se observa que este retardo es constante y dependiente del valor de p , es decir de la tasa de bit generada por la fuente.

El retardo máximo sería para $p = 1$ con una tasa de bit de 64 kbps ($\cong 5.86$ ms)

El retardo mínimo sería para $p = 30$ con tasa de bit de 2 Mbps ($\cong 0.195$ ms)

En la figura 30 se observa el comportamiento del retardo de formación de la celda a medida que se utiliza una tasa de bit mayor. Para esta gráfica, se calcula el retardo al considerar el intervalo de $p = 1$ hasta $p = 30$ (64 kbps - 2 Mbps).

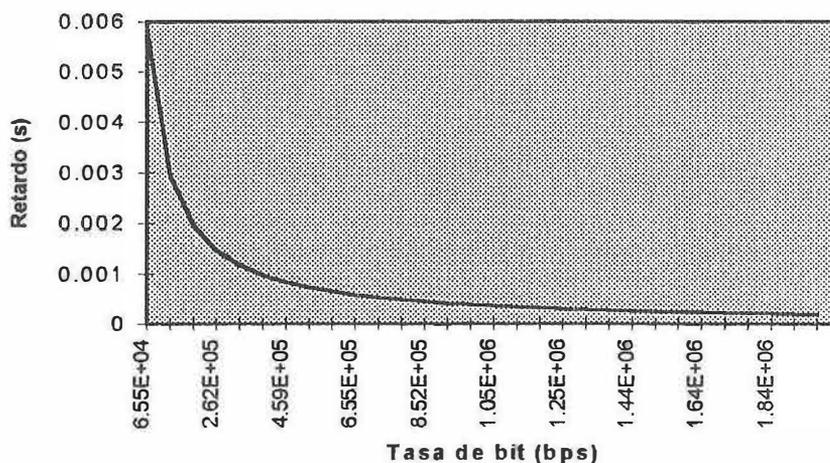


Figura 30. Variación del retardo de paquetización en función de la tasa de bit.

De la figura 30 se obtiene que a medida que la tasa de bit utilizada sea más alta, el retardo de paquetización producido tenderá a ser más pequeño. También se observa que el retardo de paquetización máximo (6 ms) es apenas una pequeña fracción del retardo total permisible por el servicio (100-200 ms), y como ya de mencionó, disminuirá significativamente a medida que se utilice una tasa de bit más alta. De hecho, el retardo variable máximo que puede producirse en el codificador a tasa de bit constante (250 ms) es de mucha mayor magnitud que el retardo máximo de paquetización.

VI.1.2 Fuente de video a tasa de bit variable

En este análisis se supone lo siguiente:

- 1) El ensamblador de paquetes es capaz de discriminar entre inicio y fin de la información codificada correspondiente a un determinado cuadro.
- 2) Por lo anterior, el ensamblador no puede mezclar información de diferentes cuadros en una celda.
- 3) La tasa de bit utilizada para transmitir la información de un cierto cuadro es constante en el intervalo de tiempo correspondiente (la cantidad de bits codificados por un cuadro se transmite a una cierta tasa de bit constante). Sin embargo, la tasa de bit varía en la transmisión de otros cuadros, dependiendo de la cantidad de bits codificados por cuadro.

En este caso, el tiempo de espera de datos para formar una celda no sería constante para todos los cuadros, por lo que el retardo de formación de la celda dependería de la tasa de bit utilizada para transmitir cada cuadro. Se define a este retardo como $R_{p,n}$, para indicar el retardo de paquetización de la tasa de bit variable a la cual se transmite un ene-

ésimo cuadro codificado. Este retardo será directamente proporcional al tamaño de la celda, e inversamente proporcional a la tasa de bit a la cual se transmite un determinado cuadro, de acuerdo a la siguiente expresión:

$$R_{p,n} = \frac{384}{x(n)} \quad \left[\frac{\text{seg}}{\text{celda}} \right] \quad (25)$$

donde $x(n)$ es la tasa de bit utilizada para transmitir el n-ésimo cuadro. Esta tasa de bit es variable, así que el retardo de paquetización consecuentemente es variable en este caso.

La figura 31 muestra la tasa de bit $x(n)$ expresada en bits por segundo de la secuencia "salesman". Se observa la similitud entre ésta y la figura 12, lo que demuestra la equivalencia de expresar la tasa de bit ya sea en bits por cuadro o en bits por segundo. El valor promedio resultó de 720 kbps, el valor pico de 730 kbps, y el valor mínimo 701 kbps.

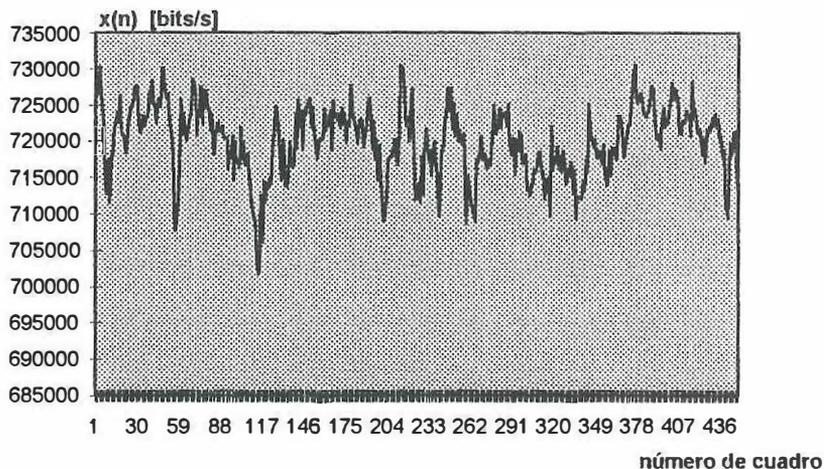


Figura 31. Tasa de bit de la secuencia "salesman" expresada en bits por segundo.

En la figura 32 se observa el comportamiento temporal del retardo de paquetización obtenido al analizar la tasa de bit generada por la secuencia “salesman”. El retardo promedio resultó de 0.533 ms, el retardo máximo de 0.547 ms y el mínimo de 0.526 ms. El retardo es bastante bajo debido a la alta tasa de bit generada por el procedimiento de codificación empleado.

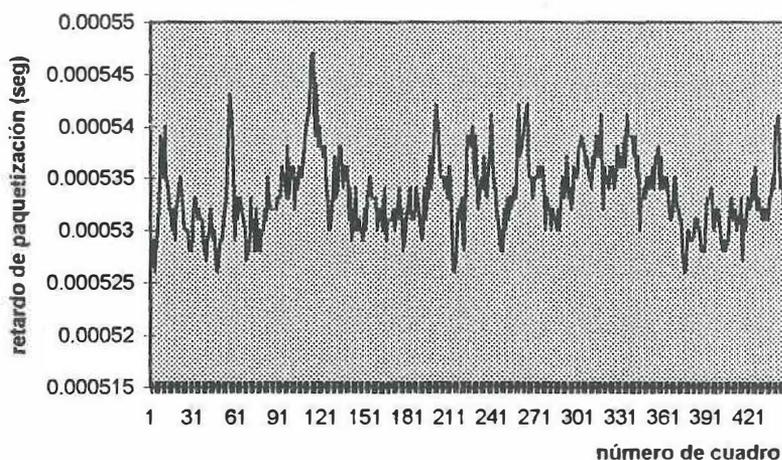


Figura 32. Retardo de paquetización a tasa variable para la secuencia “salesman”.

De la figura 32 puede observarse el retardo de paquetización, como se esperaba resulta variable. Pero los valores observables indican que el retardo es demasiado pequeño comparado con el retardo de almacenamiento temporal en el nodo de acceso. El valor máximo de paquetización resultó de 550 μ s y correspondió a la transmisión con tasa de bit más baja de la secuencia, y si se compara a los más de 100 ms permisibles para la transmisión dentro de la red, se concluye que el retardo de paquetización no contribuye en gran medida con este retardo.

La razón que más influye en los bajos retardos de paquetización son las altas tasas de bit utilizadas en la transmisión de video a tasa de bit variable. Por ejemplo, la mínima tasa de bit utilizada para la secuencia "salesman" fue de aproximadamente 700 kbps, mientras que para H.261 la mínima tasa de bit utilizada fue 64 kbps, más de 10 veces más alta en el caso de tasa de bit variable.

Además se debe considerar que los experimentos se realizaron con la calidad de imagen de menor resolución (QCIF). Se espera que al utilizar más alta resolución (CIF, TV comercial, HDTV), la tasa de bit se incremente de manera considerable, por lo que el retardo de paquetización se reduce drásticamente.

Se ha visto que el retardo de paquetización depende fuertemente de la tasa de llegada de la información de video al ensamblador de paquetes. Esto significa que las aplicaciones de video a baja tasa de bit constante (64 -128 kbps) deben de considerar el efecto de la paquetización como proceso de generación de retardos. A medida que la tasa de bit se incrementa (aplicaciones que demanden mayor ancho de banda), el retardo de paquetización empieza a disminuir considerablemente. Aunque este retardo puede resultar variable, la tasa promedio de transmisión utilizada en tales aplicaciones es lo suficientemente alta como para producir un retardo de paquetización que no contribuya prácticamente en el retardo total producido en la red. Esto significa, de acuerdo a la expresión (6) del capítulo IV, que el parámetro R_p tenderá a ser bastante pequeño y que la mayor parte del retardo dentro de la red de conmutación de paquetes se llevará a cabo en el almacenador temporal del nodo de acceso. Esto lleva a concluir que para mantener la

interactividad del servicio es suficiente que el retardo en el nodo de acceso sea menor a 100 ms, es decir:

$$R_t \leq 100 \text{ ms} \quad (26)$$

lo que limita el tamaño del almacenador temporal del nodo de acceso hasta 100 ms máximo.

VI.2 Análisis de retardos de almacenamiento temporal

Una vez definido el modelo para la tasa de bit total producido por N fuentes independientes de video, a continuación se presenta el modelo de la cola presente en el almacenador temporal común del nodo de acceso a la red de conmutación de paquetes. El modelo para el nodo de acceso a la red de conmutación de paquetes se presenta en la figura 33.

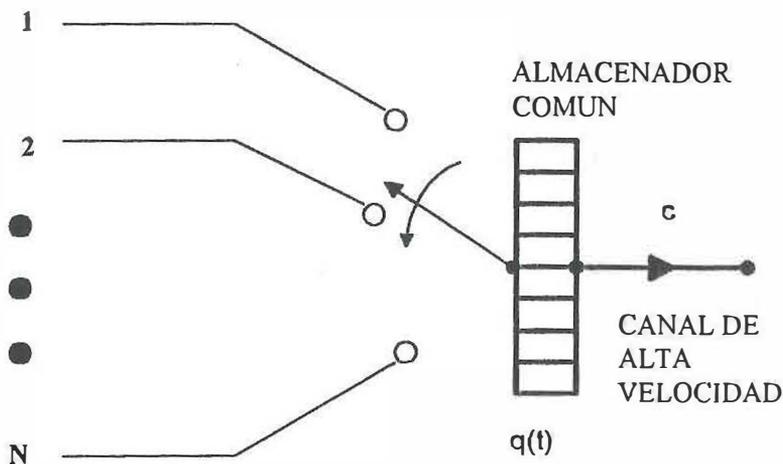


Figura 33. Modelo para el nodo de acceso.

En este modelo, N fuentes independientes de señales de video a tasa variable ya paquetizadas en el EDP, son multicanalizadas y acumuladas en el almacenador principal. Los paquetes se acumulan en el almacenador principal (común) y son enviados al canal de alta velocidad mediante un orden primero que entra-primero que sale. El almacenador temporal está siendo vaciado a una velocidad de c bits por segundo, denominado la tasa de servicio del nodo o velocidad del canal de alta velocidad. Dependiendo de la tasa total instantánea agregada de las N fuentes de video y de la tasa de servicio c , dentro del almacenador se forma una cola de espera $q(t)$.

El enfoque principal de este análisis de multicanalización estadística de señales de video, es el de llegar a una estimación de la distribución del retardo de paquetes debido a la multicanalización. Sin embargo, aunado a este análisis se deben de considerar otros aspectos de la multicanalización. Es decir, en base a la estimación del retardo, para un número dado de señales de video (cuyas características son conocidas), se puede ser capaz de estimar la capacidad de transmisión que debería ser asignada a la salida del multicanalizador, la utilización resultante del canal y el tamaño del almacenador común. Se pueden producir estimaciones para la probabilidad de sobreflujo del almacenador y la tasa de descarte de celdas.

Este modelo ha sido utilizado exitosamente en otros estudios para transmisión de datos [Anick, *et al.*, 1982] y también para el caso de video a tasa de bit variable [Maglaris, *et al.*, 1988], donde se obtiene una solución para las estadísticas de la cola de estado continuo (equivalente al flujo de fluido), considerando como entrada al modelo de flujo de fluido el modelo binomial desarrollado para la tasa de bit total.

En Maglaris *et al.* [1988] se utilizó el modelo de flujo de fluido, porque es más simple de implementar y produce resultados consistentes con simulaciones realizadas con modelos de flujo discreto (paquetes). Como resultado de tales simulaciones, se demuestra la congruencia entre los resultados de la simulación con paquetes discretos y el análisis de flujo de fluido.

En el modelo, una cola de estado continuo es alimentada por una fuente con tasa $x_N(t)$ unidades de flujo (bits/s) como entrada. La tasa del flujo de entrada puede asumir niveles discretos $(0, A, 2A, \dots, MA)$ conforme al modelo desarrollado para la tasa de bit total. Las tasas de transición exponencial del nivel i al nivel j son denotadas por $r_{i,j}$. La cola se vacía con una tasa fija de c unidades de flujo por unidad de tiempo.

La cantidad instantánea de datos presente en el almacenador temporal, es la diferencia entre la tasa de servicio c y la tasa instantánea total de bit en la entrada al multicanalizador $(iA - c)$, es decir, el sistema almacena la información entrante que está en exceso de la máxima tasa de transmisión, c , del canal de salida de alta velocidad. Una vez que el almacenador esté vacío, permanece así mientras $iA < c$. Por lo tanto, se presentan dos casos de análisis, uno es la transmisión sin retardos de almacenamiento ($iA < c$), y el otro es el más común y de mayor interés, que se presenta cuando existen retardos en el almacenador ($iA > c$).

VI.2.1 Transmisión sin retardos de almacenamiento

De acuerdo al modelo para la tasa total de bit, ésta producirá un máximo de información instantánea equivalente a MA bits/s. Si esta tasa de bit cumple que $MA < c$,

entonces no se producirán retardos de almacenamiento en ningún momento (el almacenador estará siempre vacío), ya que la probabilidad de que la tasa total de bit sea mayor a MA bits/s, es cero. Aquí se desea obtener una estimación de la máxima cantidad N de fuentes que pueden ser multicanalizadas sin que se produzcan retardos de almacenamiento, o en su caso, la mínima tasa de servicio c que debe ser asignada en el canal de alta velocidad para un determinado número N de fuentes de video. La cuestión principal radica en cómo relacionar la cantidad MA con el número N de fuentes que producen la tasa total de bit. De acuerdo a la ecuación (22) del capítulo V, el nivel de cuantización A está parametrizada por el número máximo de estados M y por N . Sin embargo, en [Maglaris, *et al.*, 1988] y [Ohta, 1994] llegan a la conclusión de que es suficiente considerar 20 niveles de cuantización por fuente, i.e., $M=20 \times N$ (produciéndose así un nivel de cuantización constante e independiente del número de fuentes). Por lo tanto, la capacidad mínima de servicio c que se debe asignar para multicanalizar N fuentes sin producirse retardos de almacenamiento resulta de:

$$c > 20 \times N \times A \quad \text{ó} \quad N < c / (20 \times A) \quad (27)$$

donde N significa el máximo número de fuentes que pueden ser multicanalizadas sin producir retardo de almacenamiento, dado una cierta tasa de servicio c .

En la expresión (27), A se obtiene de (22) en el capítulo anterior considerando que M está parametrizada por N , tal y como se definió anteriormente.

En la figura 34 se observa el comportamiento obtenido de (27) de la capacidad de servicio mínimo que debe ser asignada a la salida del nodo de acceso, para que no se produzcan retardos, para N fuentes independientes.

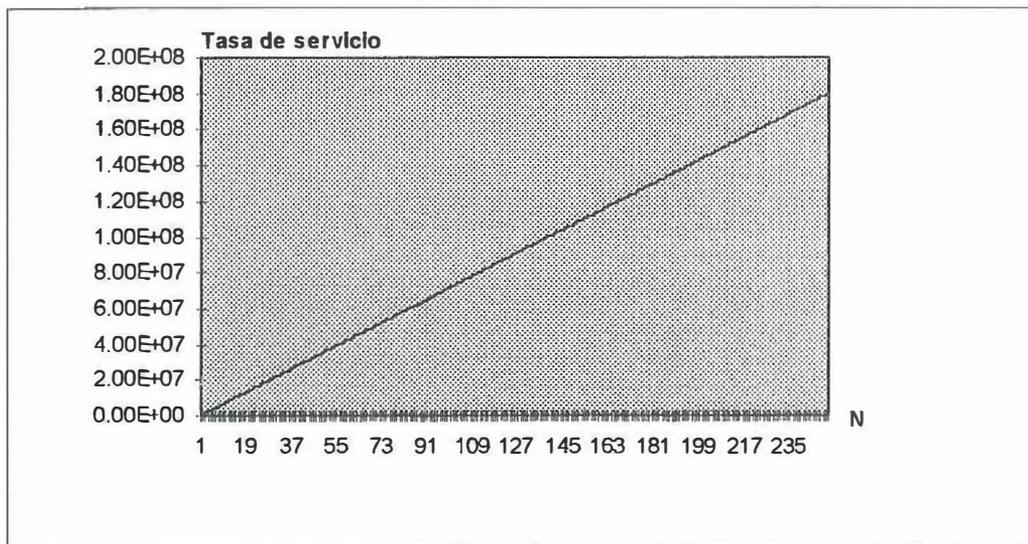


Figura 34. Tasa de servicio c mínima para no producir retardos de almacenamiento.

La estimación dada por (27) puede resultar útil para realizar una estimación preliminar de c . Al usarla, se está considerando la tasa total máxima posible del modelo. Sin embargo, generalmente sucede que este estado tiene una probabilidad de ocurrencia muy baja, por lo cual la estimación estará sobrada en capacidad y resultará en una subutilización del canal, i.e., desperdicio del ancho de banda, que es precisamente lo que se desea evitar al máximo en un buen diseño en telecomunicaciones.

El análisis anterior tienen la ventaja de no presentar el problemas de retardo en el almacenador y, por lo tanto, tampoco de sobreflujo ya que supone que la cantidad de datos que arriba instantáneamente al almacenador no sobrepasa a la tasa de servicio del

nodo. Debido a esta razón, no existirán pérdidas de información por descarte y la degradación del servicio tiende a reducirse por este concepto. Pero la subutilización del canal es un factor ponderante que obliga a buscar alternativas para el diseño eficiente del sistema, además de que en las redes de conmutación de paquetes prácticas es común el retardo por almacenamiento temporal y una probabilidad de pérdida de paquetes por sobreflujo del almacenador diferentes a cero. Otra cuestión importante, resulta de la característica del servicio de video de permitir cierto retardo máximo de transmisión, e inclusive, soporta cierta cantidad de pérdidas de información sin que esto degrade significativamente el servicio. Por estos motivos es conveniente realizar el análisis para el caso en el cual el retardo de almacenamiento temporal es mayor a cero, con la finalidad de determinar la distribución de la pérdida de paquetes y basarse en esto para obtener los valores óptimos de operación de la red.

VI.1.2 Transmisión con retardo de almacenamiento

Para este caso se considera que la tasa de bit total instantánea puede cumplir $(iA - c) > 0$. Se define el tamaño de la cola como $q(t)$ (la información instantánea presente en el almacenador). Sea la probabilidad de que la información acumulada en el almacenador en el tiempo t no exceda el valor de x (el tamaño del almacenador):

$$P_i(t, x) = P\{x_N(t) = iA, q(t) \leq x\}$$

Se define la utilización ρ como

$$\rho = \frac{E[x_N(t)]}{c} < 1 \quad (28)$$

Si $\rho \geq 1$ entonces la tasa promedio total de entrada superaría la tasa de servicio del sistema y significa que la cola presente en el almacenador crecería cada vez más y el sistema se volvería inestable. Aplicando (7) del capítulo V en la expresión (28) y resolviendo para N resulta:

$$N < \frac{c}{E[x(t)]} \quad (29)$$

que es el número máximo de fuentes con las cuales se puede cargar al sistema sin que se produzcan inestabilidades.

Si $F(x)$ representa el valor en estado estable de $P(t,x)$, la siguiente ecuación diferencial es obtenida:

$$(iA - c) \frac{dF_i(x)}{dx} = \sum_{j \neq i} r_{i,j} F_j(x) - F_i(x) \sum_{j \neq i} r_{i,j}, \quad 0 \leq i \leq M$$

con las condiciones iniciales

$$F_i(x) = 0 \quad x < 0$$

$$F_i(0) = 0 \quad iA > c \quad (30)$$

y la condición adicional cuando $x \rightarrow \infty$

$$F_i(\infty) = P\{x(t) = iA\}$$

El conjunto de ecuaciones (30) se escribe en forma matricial al usar el vector

$$F(x) = (F_0(x), \dots, F_M(x))^T$$

$$DdF(x) = RF(x) \tag{31}$$

donde D es una matriz diagonal con los elementos $(iA-c)$ y R es la matriz de transición de estados obtenida de (30). La solución para (31) está dada en términos de su valor conocido en $x = \infty$, los eigenvalores y correspondientes eigenvectores de $D^{-1}R$ como sigue:

$$F(x) = F(\infty) + \sum_i k_i \Phi_i e^{z_i x} \tag{32}$$

donde Φ_i y z_i denotan los eigenvectores y eigenvalores de $D^{-1}R$. La suma en (32) se toma sobre todos los eigenvalores en el semiplano izquierdo para que la solución sea una distribución de probabilidad en estado estable. Las constantes k_i en (32) son determinadas por las condiciones iniciales

$$F_j(0) = 0 = F_j(\infty) + \sum_i k_i \phi_{ij} \quad \text{para } c/A < j \leq M$$

La distribución en estado estable de tamaño del almacenador x está dado por

$$F(x) = P\{q(t) \leq x\} = \sum_{i=0}^M F_i(x) \quad (33)$$

y la probabilidad de que el tamaño del almacenador exceda una cierta longitud, es:

$$\bar{F}(x) = 1 - F(x) \quad (34)$$

La expresión $\bar{F}(x) = P\{q(t) > x\}$ en estado estable representa la fracción de datos que comparten el multicanalizador cuando su tamaño de cola excede un umbral x . Para video en tiempo real, hay un tamaño de cola máximo disponible x_0 (un máximo de tiempo de espera en el almacenador del nodo de acceso); cualquier paquete que comparta una cola con más de x_0 datos llegará a su destino demasiado tarde para decodificar la secuencia de imágenes de la cual forma parte. De esta forma, la función $\bar{F}(x_0)$ puede ser utilizada como probabilidad de pérdidas.

Para el caso especial del modelo de entrada, los coeficientes de transición en (9) simplifican (30) a:

$$-c dF_0(x)/dx = -M\alpha F_0(x) + \beta F_1(x), \quad i = 0$$

$$(iA-c) dF_i(x)/dx = (M-i+1)\alpha F_{i-1}(x) + (i+1)\beta F_{i+1}(x) - [i\beta + (M-i)\alpha] F_i(x), \quad 0 < i < M$$

$$(MA-c) dF_M(x)/dx = \alpha F_{M-1}(x) - M\beta F_M(x), \quad i = M$$

El procedimiento para la resolución del sistema de ecuaciones se codificó y los resultados se presentan a continuación en el análisis de resultados. Se utilizaron los parámetros entregados por el modelo de la tasa de bit total obtenido en el capítulo V.

VI.3 Análisis de resultados

En la figura 35 se presentan las curvas de distribución de las pérdidas para diferentes valores de niveles de cuantización M considerando una sola fuente de video. Estas curvas indican la sensibilidad del modelo al número de niveles de cuantización. La finalidad es encontrar el número adecuado para M que será utilizado en el análisis de la distribución de pérdidas para el caso de múltiples fuentes.

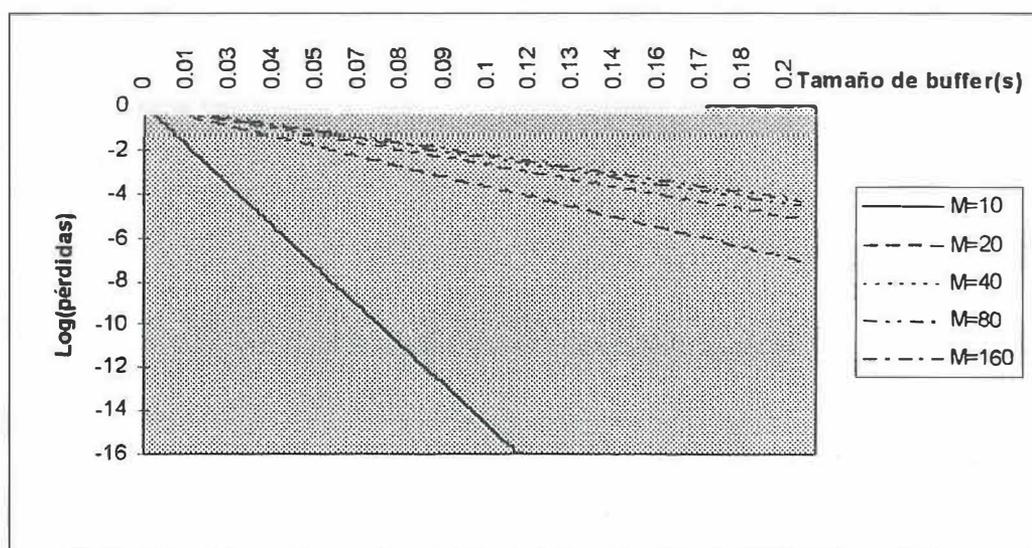


Figura 35. Comportamiento del modelo para diversos valores de M . $N = 1$, $\rho = 0.9996$, $C = 720288$ bps.

Cabe recordar que entre más grande sea el valor de M , mejor será la aproximación del modelo. Por cuestiones de limitantes de cálculo en la programática utilizada, el número máximo de niveles de cuantización utilizado fue de 160, por lo que con este valor se

obtienen las curvas más confiables, y se toma como referencia para las curvas obtenidas con M menor.

Es notoria la diferencia existente entre los resultados para $M=10$ y $M=160$, ésto significa que los resultados para $M=10$ no serían confiables. Para $M=20$ la cercanía entre las curvas resulta bastante significativa. La tendencia de las curvas es separarse entre sí conforme el intervalo de cálculo es más grande, es decir, conforme el valor utilizado para el tamaño de almacenador crezca. Sin embargo, debido a las restricciones impuestas al tamaño del almacenador (100 ms máximo), la diferencia entre las curvas para $M=20$ y $M=160$ es mínima para el intervalo de interés de 0-100 ms.

En los casos de $M=40$ y $M=80$, se observa la tendencia de que las curvas se acercan aún más a la curva de referencia, lo que corrobora la suposición de que mientras M sea más grande, la precisión resultante será mejor.

Otra característica observable en las curvas es que el considerar una M no adecuada puede resultar en subestimar la probabilidad de pérdidas en el almacenador lo que significa que se estima una probabilidad de pérdidas que puede estar muy por debajo de las pérdidas reales. Esto se obtiene al observar que las curvas con M más pequeña se aproximan "por debajo" de la curva de referencia. Tal subestimación resulta mínima si se encuentra un valor adecuado para M .

De acuerdo al análisis anterior, lo más adecuado es seleccionar $M=160$ para obtener valores de pérdidas confiables. Por otra parte, es conveniente definir un nivel de cuantización A constante para todos los casos de estudio y experimentos que se realicen. De acuerdo a la expresión (22) del capítulo V, esto se logra al parametrizar el número de

niveles de cuantización en función de la cantidad de fuentes multicanalizadas, lo que resultaría en proponer $M=160 \times N$. Sin embargo, esto restringe los casos de estudio a una sola fuente de video ya que se tiene capacidad limitada en la programática donde se lleva a cabo el análisis, como se observó en el capítulo V. Por lo tanto, se sacrifica la precisión de los resultados en aras de obtener el comportamiento más general del sistema cuando se incrementa el número de fuentes. De acuerdo a la discusión anterior si se utiliza M entre 20 y 80, la precisión resultante en el intervalo deseado es bastante aceptable. Por lo tanto, si se define $M=20 \times N$, la cantidad máxima de fuentes a multicanalizar es mayor comparado con los otros casos, así que se eligió este valor, mismo que se utiliza en el modelaje de la tasa de bit agregada del capítulo V y corrobora la validez de ese resultado.

Esto concuerda con los estudios presentados por Maglaris *et al.* [1988] y Ohta [1994], que concluyen que los datos generados en sus experimentos pueden ser modelados con suficiente precisión al utilizar $M=20 \times N$. También esos estudios respaldan sus conclusiones con simulaciones del sistema de multicanalización hechas al respecto, que demuestran la concordancia entre los resultados de análisis y simulación.

La figura 36 muestra comparativamente las curvas obtenidas en la distribución de las pérdidas de datos al simular la multicanalización hasta de 8 fuentes de video, y variar el tamaño del almacenador entre 0-200 ms. La utilización se fijó en 0.9996. El número de niveles de cuantización fue $M=20 \times N$. Una observación importante es que valores de la utilización menores a 0.9996 produjeron probabilidades de pérdidas igual a cero.

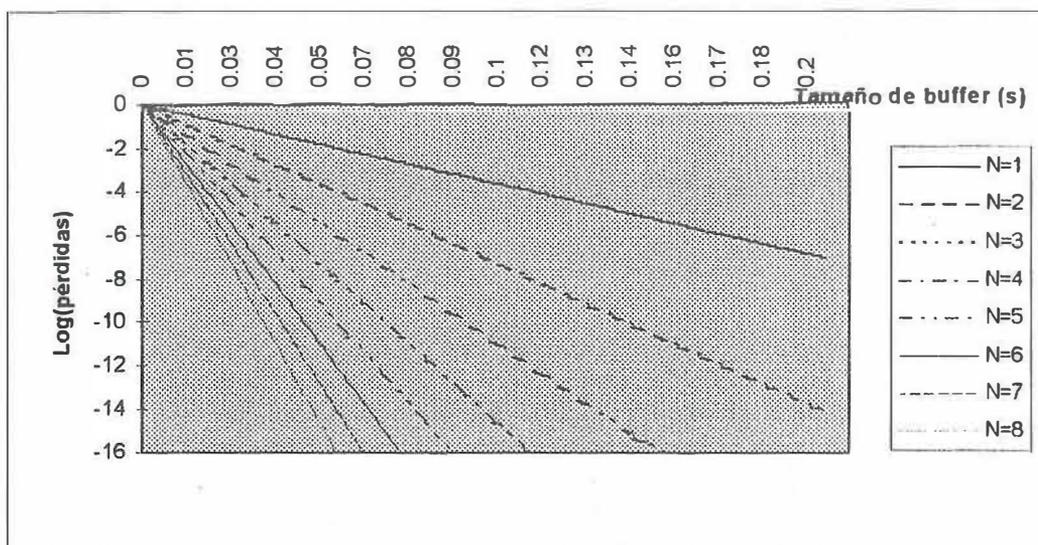


Figura 36. Comportamiento de las pérdidas para el caso de múltiples fuentes.
 $\rho=0.9996$.

De la figura 36 se obtiene el efecto de la multicanalización estadística, ya que al aumentar el número de fuentes (y bajo una misma utilización) las probabilidades de pérdidas disminuyen considerablemente. Debido a limitaciones en la programación utilizada, la mínima probabilidad de pérdidas que se pudo obtener fue del orden de 10^{-16} , por lo que la gráfica se corta en este valor. El hecho de que la utilización se mantenga constante al aumentar el número de fuentes implica que la tasa de servicio del canal de alta velocidad se incrementa de acuerdo al número de fuentes siendo multicanalizadas y a la expresión (28).

Otro de los aspectos que puede observarse en la misma figura, es que al aumentar el tamaño del almacenador se producen menores pérdidas por rebalse. Esto resulta lógico ya que un almacenador más grande tiene menor probabilidad de que se desborde que uno más pequeño, cuando la cantidad de paquetes que entra a éstos es la misma. Pero no se puede aumentar arbitrariamente el tamaño del almacenador ya que esto propiciaría mayor

retardo de permanencia en el mismo, lo cual pudiera afectar la interactividad del servicio de video. Como resultado del análisis del capítulo IV el tamaño máximo del almacenador temporal es de 100 ms.

Por otro lado, la calidad del servicio se deteriora si la red produce pérdidas excesivas de paquetes. Por lo tanto, es necesario que la probabilidad de pérdidas se mantenga lo más baja posible o al menos debajo de un valor preestablecido. La obtención directa de este valor está fuera del alcance de este trabajo, es decir, la determinación de como se afecta la calidad de la imagen en función de la tasa de descarte de paquetes en el nodo de acceso. Sin embargo, se puede suponer que el servicio de video (para fines de análisis) puede soportar un máximo de pérdidas de 10^{-6} .

De la figura 36, puede obtenerse información adicional para seleccionar el valor óptimo para el tamaño del almacenador y la tasa de servicio que debe asignarse al nodo de acceso con la finalidad de dar servicio hasta 8 fuentes de video. Esta información está relacionada con el tamaño de almacenador temporal que se requiere a medida que se agreguen más fuentes al nodo de acceso, así como la probabilidad de pérdidas que se generan con un cierto tamaño de almacenador también cuando se agregan fuentes de video al nodo.

La figura 37 presenta el comportamiento del tamaño del almacenador que asegure una cierta probabilidad de pérdidas máxima en función del número de fuentes conectadas al nodo.

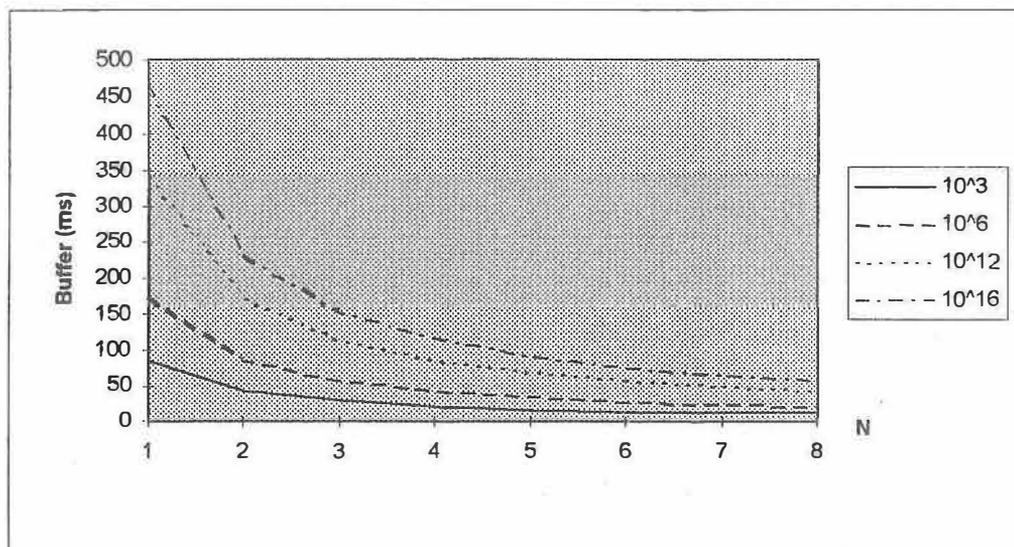


Figura 37. Requerimientos en el tamaño del almacenador en función del número de fuentes multicanalizadas y de la probabilidad de pérdidas que genere.

Lo primero que se observa en la figura 37 es la reducción del tamaño del almacenador temporal a medida que se multicanalicen un mayor número de fuentes. También se obtiene que la tendencia de las curvas es a estabilizarse en un cierto valor a medida que el número de fuentes aumenta. Esta tendencia, en el caso de que siempre estuvieran siendo multicanalizadas un número grande de fuentes de video, indicaría el tamaño óptimo del almacenador que asegurara una probabilidad máxima de pérdidas. Pero si el número de fuentes multicanalizadas es bajo, esa tendencia no se presenta en las curvas. Se observa también, que las curvas tienden a estabilizarse con un menor número de fuentes de video mientras la probabilidad de pérdidas sea más alta. Por ejemplo, si se multicanalizaran 6 ó más fuentes de video, el tamaño de almacenador óptimo que asegura una máxima probabilidad de pérdidas de 10^{-6} es de aproximadamente 20 ms, de acuerdo a la figura 37.

La tasa de servicio requerida depende del número de fuentes, y se obtiene mediante la expresión (28).

Sin embargo, este estudio considera un número bajo de fuentes a multicanalizar, por lo que el método anterior no es aplicable completamente. El problema ahora radica en optimizar la transmisión para el caso de una fuente de video.

Para esto, la figura 38 presenta el comportamiento de las pérdidas en función del número de fuentes y bajo algunos valores de almacenador seleccionados. Los datos utilizados para generar la figura 38, se extraen de la figura 36.

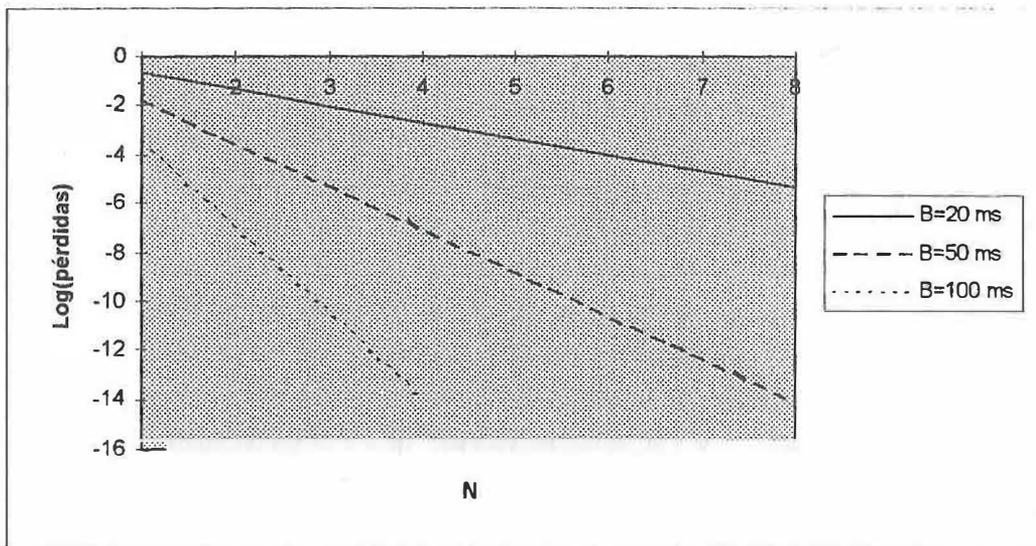


Figura 38. Comportamiento de las pérdidas en función del número de fuentes y con un determinado tamaño de almacenador.

Si se utiliza el tamaño de almacenador temporal de 20 ms seleccionado para el caso de muchas fuentes, se aseguraría por un lado que el retardo de permanencia en el almacenador fuera menor a 100 ms, como se desea. Por otro lado, la probabilidad de

pérdidas crece hasta 10^{-1} , de acuerdo a la figura 34, tomando como referencia a la curva de $B=20$ ms, y la probabilidad de pérdidas requerida es 10^{-6} . Por lo tanto, este tamaño de almacenador generaría demasiados paquetes perdidos para el caso de que estuviera siendo atendida únicamente una fuente en el nodo. De la figura 37 también se obtiene que el tamaño mínimo de almacenador requerido en el caso de una sola fuente de video es de 180 ms aproximadamente, para asegurar una probabilidad de pérdidas menor a 10^{-6} . Sin embargo, esto generaría un retardo de almacenamiento bastante superior al permitido por el servicio de video. Para probabilidades de pérdidas menores, se requiere que el almacenador sea más grande. Por ejemplo, para una probabilidad máxima de pérdidas de 10^{-12} , se requiere un almacenador cercano a los 350 ms, inaceptable para el servicio.

De la figura 38, para $B=100$ ms la curva muestra que para los casos de 2 hasta 8 fuentes, un almacenador de 100 ms produce la probabilidad de pérdidas máxima deseada. Por ejemplo, para 8 fuentes un almacenador de 100 ms de tamaño produciría una tasa de pérdidas mucho menor a 10^{-16} , de acuerdo a la curva de 100 ms de la figura 38 y para 2 fuentes de aproximadamente 10^{-7} . Esto significa que, garantizando que se cumplan las condiciones requeridas de retardos y pérdidas para el caso de una fuente de video, éstas también se cumplen para un número mayor de fuentes.

De acuerdo al análisis presentado, la condición para una fuente de video no puede ser alcanzada. La alternativa a seguir es disminuir la utilización del canal de alta velocidad, es decir, aumentar la tasa de servicio del canal con la finalidad de disminuir las pérdidas y producir la condición de operación deseada. Sin embargo, la velocidad del canal no se

debe aumentar indiscriminadamente ya que la utilización podría reducirse drásticamente lo que equivaldría a desperdiciar el ancho de banda del canal. Las gráficas presentadas fueron obtenidas con datos generados al forzarse una utilización del canal bastante alta de 0.9996. Los experimentos demostraron que al reducir la utilización unas cuantas décimas, por ejemplo a 0.99 o más abajo, el modelo producía una probabilidad de pérdidas de cero. En este caso, reducir la utilización más abajo de 0.99 significa incrementar la tasa de servicio innecesariamente, ya que la probabilidad de pérdidas se mantiene, obviamente, en cero, y se estaría desperdiciando el canal de alta velocidad.

Por lo tanto, para el caso de estudio que se presentó, el tamaño de almacenador óptimo es de 100 ms y la utilización más conveniente podríamos asignarla de 0.99, lo cual equivaldría a una tasa de servicio de aproximadamente 5.82 Mbps y el tamaño del almacenador de 582 kbits.

La razón por la cual el modelo se comporta de esta manera proviene de las características de la tasa de bit variable producida por los experimentos presentados en el capítulo III. Especialmente se debe a la baja relación entre la desviación estándar y la media de los datos. Esta característica impidió obtener un comportamiento más general de la distribución de las pérdidas en función de la utilización del canal, el cual es bastante importante al intentar asignar la tasa de servicio óptima.

Podrían hacerse otras consideraciones con respecto al tamaño del almacenador. Por ejemplo, podría ser que se deseara un retardo máximo de 20 ms en el nodo de acceso, y de nuevo se requiere una pérdida máxima de 10^{-6} . De acuerdo a la figura 38 y la curva B=20 ms, para ningún caso la probabilidad de pérdidas es menor a 10^{-4} , por lo que la única

opción viable es modificar de nueva cuenta a la tasa de servicio. De igual manera que en el caso anterior se elige una utilización del canal de 0.99 para producir que la probabilidad de pérdidas caiga a cero. Así, la tasa de servicio es nuevamente de 5.82 Mbps y el tamaño del almacenador de 116.4 kbits.

En conclusión, se puede hacer notar que de la totalidad de los parámetros que se involucran en el modelaje del nodo de acceso, el tamaño del almacenador y la velocidad del canal de alta velocidad son los que finalmente son considerados para determinar completamente el diseño del nodo de acceso a la red de conmutación de paquetes, de acuerdo a las necesidades de transmisión del servicio de videoconferencia.

Para el caso de estudio tratado, en la tabla V se muestran los resultados del análisis para transmitir hasta 8 fuentes de videoconferencia, que permiten al retardo extremo a extremo estar dentro de los límites de interactividad.

Tabla V. Valores numéricos de los parámetros operativos para videoconferencia

Número máximo de fuentes de video	8 fuentes
Retardo extremo a extremo	Menor a 200 ms
Tamaño de almacenador	582 kbits
Velocidad del canal de alta velocidad	5.82 Mbps
Tasa de pérdidas de paquetes	10^{-6}

En este capítulo se presentaron los análisis de retardos que se originan dentro de la red de conmutación de paquetes. El retardo se origina en dos procesos: la paquetización y el almacenamiento temporal en el nodo de acceso. El análisis de retardos demostró que el retardo permisible en la red se produce en el almacenador temporal del nodo de acceso y que se encuentra limitado a un valor máximo de 100 ms. Este valor a su vez limita el tamaño del almacenador máximo, que garantiza la interactividad del servicio.

Por otra parte, la red de conmutación de paquetes debe garantizar, además de retardos limitados, un número limitado de paquetes perdidos con la finalidad de no degradar el servicio. Estos dos factores son clave para la obtención de valores óptimos para el tamaño de almacenador y la tasa de servicio en el canal de alta velocidad.

Para el caso de estudio de transmisión de 8 fuentes de videoconferencia máximo a tasa de bit variable, el tamaño recomendado del almacenador resultó de 582 kbits, el canal de alta velocidad de 5.82 Mbps, y la probabilidad de pérdidas en el nodo de acceso por sobreflujo es menor a 10^{-6} .

VII CONCLUSIONES Y RECOMENDACIONES

VII.1 Conclusiones

El estudio ha considerado innovativamente la transmisión de la señal de videoconferencia a través de redes de conmutación de paquetes (normalmente se realiza a través de redes de conmutación de circuitos). De acuerdo a las características de la señal de videoconferencia se desprende que este tipo de servicios, a diferencia de los datos, genera un flujo de datos continuo, que debe ser entregado con restricciones de tiempo determinadas por las características de interactividad del servicio. La necesaria digitalización de las señales de video genera una gran cantidad de información binaria que debe ser reducida con técnicas de compresión de video. Las técnicas de compresión de video más comunes en la actualidad producen inherentemente una cantidad de información variante que dependerá del movimiento presente en la secuencia de video. La adaptación de estas técnicas a sistemas de comunicaciones basados en esquemas de conmutación de circuitos, tal como lo es la red digital de servicios integrados RDSI, ha propiciado la aparición del estándar ITU-T H.261, el cual genera una tasa de bit constante a costa de variar la calidad del servicio para mantener así el ancho de banda permisible en este tipo de redes. Se genera, de esta manera, un compromiso entre la tasa de bit (dependiente de la tasa de compresión requerida) y la calidad del servicio. Una característica importante de la información producida por éste estándar es su sensibilidad a los retardos producidos en la transmisión.

Por otra parte, la red de conmutación de paquetes realiza procesos que producen retardos tanto fijos como variables a la información que es transmitida sobre la misma,

retardos que son producidos a causa del transporte de la señal (retardos fijos), y por efecto de almacenamiento temporal en distintos puntos de la red (retardos variables). Estos retardos pueden afectar considerablemente la calidad del servicio de video si no son mantenidos bajo un cierto valor, mismo que depende de la interactividad del servicio.

En base a lo anterior, el trabajo realizado en el presente estudio ha permitido obtener las siguientes conclusiones:

1) Las características de los servicios interactivos permiten soportar un retardo extremo a extremo entre 100 a 200 ms. La suma de los retardos fijos más los variables deben de mantenerse menor a este intervalo para evitar la degradación del servicio.

2) Los retardos de codificación dependen de la tasa de bit deseada a la salida del codificador. Para codificación a tasa de bit constante, mediante el estándar H.261, el retardo de codificación es fuertemente afectado por el sistema de regulación de la tasa de bit, que incluye un almacenador temporal anterior a la transmisión de la información codificada. Este sistema puede producir retardos de permanencia en el almacenador hasta de 250 ms máximo (retardos variables), valor que resulta demasiado alto para las exigencias de interactividad del servicio. A éste valor aún falta agregarle los retardos que normalmente se producen en la transmisión sobre la red de conmutación de paquetes, lo que produciría un retardo total excesivo como para mantener la interactividad del servicio, por lo que la calidad del mismo se degradaría fuertemente. Además, se deben agregar las posibles pérdidas de información debidas al desbordamiento en el almacenador temporal en el nodo de acceso a la red de paquetes, que contribuirían aún más a degradar la calidad del servicio.

También se debe considerar la variación indeseable sobre la calidad de video que se produce por efecto del sistema de regulación de la tasa de bit. Este efecto es más notorio a medida que se disminuye la tasa de bit a la salida del codificador de video.

3) A pesar de estos inconvenientes de transmitir video paquetizado a tasa de bit constante, una alternativa para mejorar la calidad del servicio al transmitir con este sistema de codificación es obligar a la red de conmutación de paquetes a dar servicio preferencial a esta clase de transmisión mediante esquemas de prioridad de servicio. De esta manera, el retardo de transmisión se reduce significativamente y sólo afectaría el retardo variable de codificación, mismo que puede ser eliminado por un almacenamiento temporal adecuado en el extremo receptor.

4) La transmisión de video a tasa de bit variable es posible mediante las redes modernas de conmutación de paquetes. Esto significa que se puede prescindir del sistema de regulación de la tasa de bit en el codificador, lo cual reduce significativamente el excesivo retardo generado por ese sistema. Además, la calidad del servicio puede mantenerse a un nivel mínimo deseable. Otro de los aspectos es que permite la multicanalización estadística de las señales de video, lo que puede producir mejor utilización de la tasa de bit del canal de alta velocidad. La multicanalización estadística no es posible con la transmisión de video a tasa de bit constante ya que en este caso las señales son predecibles una de otra, por lo que las mismas no cumplen con la condición de fuentes estadísticamente independientes, ya que son completamente determinísticas.

5) La red de conmutación de paquetes contiene procesos que generan retardos variables y además una cierta probabilidad de pérdidas de paquetes por efecto de sobreflujo en el almacenador temporal dentro del nodo de acceso. Los retardos variables más importantes producidos por la red de conmutación de paquetes resultan de los procesos de formación del paquete y por el almacenamiento temporal en el nodo de acceso a la red. El retardo de paquetización depende de la tasa de bit utilizada para la transmisión, así que mientras más alta sea la tasa de bit, el retardo de paquetización disminuye dramáticamente. Para tasa de bit constante el retardo de paquetización es constante, y para tasa de bit variable el retardo de paquetización es dependiente de la velocidad instantánea a la cual se transmite un cierto cuadro de video.

6) Para determinar el tamaño óptimo del almacenador en el nodo de acceso a la red (y por lo tanto el retardo variable producido dentro del mismo), se debe tomar en cuenta el retardo máximo extremo a extremo permisible por el servicio y otros factores adicionales tales como las características de la tasa de bit generadas por las fuentes de video, el número de fuentes a las cuales puede dar servicio un determinado nodo, y la tasa de servicio del canal de alta velocidad. Si se considera que el servicio de videoconferencia permite un máximo retardo de 200 ms, entonces la red de conmutación de paquetes no debe producir una suma de retardos fijos y variables superiores a este valor. Sin embargo, en la práctica resulta conveniente que los retardos se mantengan muy por debajo de este valor con la finalidad de asegurar completamente la entrega de los paquetes dentro de los límites preestablecidos de retardos. Como el presente estudio consideró solamente el caso de una red conformada por

un nodo de acceso y un nodo de salida, el retardo variable de almacenamiento temporal se produciría en el nodo de acceso, así que su tamaño debe ser limitado al máximo retardo permisible, es decir, 100 ms.

7) El retardo por efecto de la cola en el nodo de acceso resultó ser dominante sobre el retardo de paquetización (6 ms máximo para el caso de tasa de bit constante a baja velocidad de 64 kbps y 550 μ s máximo para el caso de estudio a tasa de bit variable comparado con el intervalo máximo de 100 ms mencionado).

Sin embargo, el retardo variable producido por la red de conmutación de paquetes puede ser eliminado al introducir un almacenador temporal lo suficientemente grande en la terminal destino. Pero este mecanismo se complica con la posible pérdida de paquetes que puede producirse en el almacenador del nodo de acceso a la red. Por lo que ahora el problema radica en que la red debe de garantizar un límite en la cantidad de paquetes perdidos. Por lo tanto, resultó necesario modelar las características de la tasa de bit generada por la fuente de video a tasa variable; adecuar y aplicar un modelo para el nodo de acceso que considerara todos los parámetros mencionados, con la finalidad de encontrar la distribución de la probabilidad de pérdidas de paquetes producidos en el almacenador temporal y seleccionar el tamaño de almacenador óptimo bajo los valores de operación deseables en los parámetros que conforman al modelo del nodo de acceso.

8) Para el modelaje de la tasa de bit variable, se encontró que un proceso pseudoaleatorio no le caracteriza adecuadamente ya que produce una distribución probabilística diferente a la obtenida para la tasa de bit experimental. También la ausencia de correlación en el proceso

seudoaleatorio se diferenci6 de la alta correlaci6n entre cuadros consecutivos obtenida en el an6lisis estadístico realizado a los datos de la tasa de bit generados por la secuencia de videoconferencia. Por lo tanto, modelar al sistema mediante un an6lisis para se±ales completamente aleatorias no result6 conveniente. Se opt6 por adaptar y aplicar un modelo de Markov de nacimiento y muerte que ha sido utilizado exitosamente en diferentes estudios realizados para la tasa de bit de secuencias de videoconferencia sin cambios de escena. El modelo result6 adecuado para los datos obtenidos en el trabajo experimental, a pesar de que se utiliz6 un sistema de codificaci6n y compresi6n de video diferente a los anteriores estudios. Aunque las expresiones matemáticas que se obtienen al aplicar este modelo no presentan gran complejidad, en el c6lculo de ciertos valores pueden presentarse ciertos inconvenientes si no se posee herramientas computacionales lo suficientemente potentes para tratar con n6meros mayores a 10^{+300} , que aparecen com6nmente cuando se requiere obtener gran precisi6n en el modelaje.

9) El modelaje para el nodo de acceso a la red de conmutaci6n de paquetes demostr6 la efectividad de la multicanalizaci6n estadística de la tasa de bit variable generadas por ocho fuentes de video independientes. Tambi6n demostr6 a la intrínseca relaci6n entre el n6mero de fuentes a multicanalizar, el tama±o del almacenador, la capacidad de servicio del nodo y la probabilidad de p6rdidas por sobreflujo. Si se desean disminuir la probabilidad de p6rdidas de paquetes dado un cierto n6mero de fuentes a multicanalizar, se tienen dos alternativas: aumentar el tama±o del almacenador temporal o incrementar la tasa de servicio. El aumento del tama±o del almacenador produce mayor retardo de espera dentro del mismo, lo cual

puede resultar contraproducente para la calidad del servicio, y también un tamaño de almacenador impráctico o costoso. Por su parte, al incrementar la tasa de servicio puede resultar en un desperdicio de ancho de banda o en la utilización no eficiente del canal. Por lo tanto, existe un compromiso entre ambos parámetros que debe ser resuelto de acuerdo a las condiciones tecnológicas disponibles en los sistemas de almacenamiento temporal, a las capacidades de transporte de información de los sistemas actuales de comunicaciones, y a las limitaciones económicas prevalecientes.

VII.2 Aportaciones

Entre las principales aportaciones de este trabajo se pueden mencionar las siguientes:

- 1) Determinación de las condiciones de operación de la red de conmutación de paquetes para la transmisión óptima de videoconferencia, considerándose como un punto de inicio para otros estudios que retomen diferentes aspectos de la transmisión.
- 2) Caracterización estadística y modelaje de la tasa de bit generada por secuencias de videoconferencia sin cambios de escenas.
- 3) Modelaje del nodo de acceso a la red de conmutación de paquetes para la determinación de la distribución de las pérdidas en el almacenador temporal.
- 4) Determinación del retardo extremo a extremo sobre la señal de video en redes de conmutación de paquetes.
- 5) Obtención de los retardos que más afectan la transmisión sobre este tipo de redes.

6) Definición de intervalos de operación de retardos permisibles obtenidos de la interactividad del servicio de videoconferencia.

VII.3 Recomendaciones

-Realizar estudios equivalentes que consideren secuencias de videoconferencia de mayor duración con la finalidad de obtener medidas estadísticas más confiables. Podrían considerarse también a secuencias con mayor movimiento y con cambios de escenas, tales como secuencias de televisión comercial.

-Considerar, en otros posibles estudios relacionados, la utilización de estándares de compresión diferentes a H.261. En lo especial, que consideren a la serie de estándares MPEG, los cuales ya han empezado a ser empleadas por las aplicaciones de video en los sistemas multiservicios y tienden a dominarlas en el futuro inmediato a nivel internacional.

-De acuerdo a lo anterior, modelar las tasas de bit de generadas por secuencias de video que contengan cambios de escena y su respectiva influencia en el modelo para el nodo de acceso a la red.

-Resulta conveniente realizar estudios con la finalidad de determinar la dependencia de la calidad subjetiva del servicio de video con respecto a la tasa de pérdidas de paquetes. También de los mecanismos más apropiados para evitar la degradación excesiva de la calidad del servicio cuando se presenten pérdidas de paquetes (técnicas de corrección de errores o, en su caso, técnicas de compensación de los efectos causados por paquetes perdidos sobre la calidad del servicio).

Literatura citada

- CCITT (ITU-T). 1990. "Video codec for audiovisual services at p×64 kbit/s".
Recomendación H.261.
- Chen, Ting-Chung, Shaw-Min Lei y Ming-Ting Sun. 1993. "Continuous presence video bridging based on H.261 standard", SPIE, Vol. 2094.
- Dalgic, Ismail, William Chien y Fouad A. Tobagi. 1994. "Evaluation of 10Base-T and 100Base-T Ethernet Carrying Video, Audio and Data Traffic", IEEE Infocom, pag. 1094-1102.
- Gonzales, Rafael C. y Richard E. Woods. 1992. "Digital image processing", Addison-Wesley Publishing Co.
- Haskell, B. G. 1972. "Buffer and channel sharing by several interframe picturephone coders," Bell Syst. Tech. J., vol. 51, no. 1, paginas 261-289.
- Karlsson, Gunnar y Martin Vetterli. 1988. "Subband coding of video for packet networks", Optical Engineering, Vol. 27, No. 7.
- Karlsson, Gunnar y Martin Vetterli. 1989. "Packet video and its integration into the network architecture", IEEE JSAC, Vol. 7, No. 5.
- Lee, S. H. y L. T. Wu. 1988. "Variable rate transport in broadband packet networks", SPIE, Vol. 1001, Visual communications and image processing.
- Liou, Ming. 1991. "Overview of the p×64 kbits/s video coding standard", Communications of the ACM, Vol. 34, No. 4.

Literatura citada (continuación)

- Maglaris, B, D. Anastassiou, P. Sen, G. Karlsson y J. D. Robbins. 1988. "Performance models of statistical multiplexing in packet video communications," IEEE Trans. on Commun., vol. 36, no. 7.
- Ohta, Naohisa. 1994. "Packet video. Modeling and signal processing", Artech House.
- Okubo, Sakae. 1995. "Reference model methodology-A tool for the collaborative creation of video coding standards", Proceedings of the IEEE, Vol. 83, No. 2.
- Universidad de Stanford. 1991. "Programa de codificación y decodificación de video PVRG-P64".

APÉNDICE A

DESCRIPCIÓN DEL ESTÁNDAR ITU-T H.261

Esta recomendación describe los métodos de codificación y decodificación de video para los componentes de imágenes en movimiento de servicios audiovisuales a tasas de bit de $p \times 64$ kbit/s, donde p está en el intervalo de 1 a 30 (64 hasta 1920 kbit/s).

Este servicio se proyectó pensando en la Red Digital de Servicios Integrados (RDSI). Inicialmente el objetivo fue una recomendación de un estándar de codificación de video para transmisión en $m \times 384$ kbit/s con $m=1,2,\dots,5$. Con nuevos descubrimientos en técnicas de codificación de video, llegó a ser claro que solo un estándar ($p \times 64$ kbit/s $p=1,2,\dots, 30$), puede cubrir la capacidad del canal RDSI completa. Fué aprobada en 1990 (Okubo, 1995).

H.261 es parte de una especificación más grande llamada H.320. Esta incluye especificaciones para compresión y descompresión de la señal de audio, enmarcamiento y corrección de errores para audio, video y datos, y protocolos de señalización para el canal de comunicaciones.

H.261 está proyectado específicamente para aplicaciones en telecomunicaciones bidireccionales, ambiente en el cual es muy crítico codificar y decodificar el video en tiempo real. También es importante que sea introducido un retardo lo más pequeño posible en el flujo de video. Esto es, si el tiempo en el que una acción sucede frente a una cámara en un extremo y hasta que es visto en el otro extremo es más que una pequeña fracción de un segundo, creará una interrupción molesta en el proceso de comunicación.

Cuando $p=1$ ó 2 , debido a que la tasa de bit disponible es muy baja, solamente es recomendable para las comunicaciones visuales cara a cara (del inglés desktop face to face) como videotelefonía. Para $p \geq 6$, debido a la mayor tasa de bit disponible, imágenes más complejas pueden ser transmitidas con mejor calidad. Esto es más aplicable a videoconferencia.

Especificación general del codec H.261

A continuación se presentan brevemente las características de codificación y decodificación de la recomendación H.261 (Liou, 1991).

Entrada y salida de video

Para permitir a una sola recomendación una cobertura en y entre regiones que usan estándares de video diferentes (525-NTSC y 625-PAL líneas), el codificador opera en imágenes basadas en un formato intermedio común (CIF). Los estándares de las señales de video de entrada y salida, los cuales pudieran representarse por señales de video compuestas o mediante componentes, señales analógicas o digitales y los métodos para realizar cualquier conversión necesaria hacia y del formato de codificación fuente no están sujetos a recomendación.

También define el formato QCIF o cuarta parte de CIF. Todos los codecs deberán de operar con QCIF, el uso de CIF es opcional. La selección de CIF ó QCIF depende de la disponibilidad de la capacidad del canal. Para $p=1$ ó 2 , QCIF es usado normalmente para

aplicaciones de videotelefonía. Para $p \geq 6$, CIF podría ser usado debido a que hay más bits disponible para codificar una imagen CIF.

Los formatos CIF y QCIF están divididos en una estructura de bloques jerárquicos consistente de Imagen, Grupos de Bloques (GOB), Macro Bloques (MB) y Bloques. Cada Macrobloque está compuesto de cuatro bloques de 8×8 pixeles de luminancia (Y) y dos bloques de 8×8 pixeles de crominancia (C_B , C_R). Un Grupo de Bloques está compuesto de 3×11 Macro Bloques. Una Imagen QCIF tiene 3 GOB mientras CIF tiene cuatro veces este número.

Espacio de colores

El estandar H.261 define el espacio de colores YUV para representar a las señales de video digitales. La señal Y se define para la luminancia y las restantes U y V para la crominancia. Para representar a un elemento de imagen se emplea el formato 4:1:1, lo que significa que por cada cuatro muestras cercanas (formando un cuadro) de luminancia Y, existe una muestra de las señales de crominancia U,V.

A partir de este formato, si se considera que cada muestra es de 8 bits, y que 4 muestras de luminancia “comparten” una sola muestra de las señales de crominancia, se obtiene que la cantidad de bits por las cuatro muestras es de $4 \cdot 8 + 8 + 8 = 48$ bits. Si consideramos un promedio de las cuatro muestras, se obtiene $48/4 = 12$ bits por muestra o bits por elemento de imagen en promedio.

Algoritmo de codificación fuente

Un híbrido de predicción intercuadro para remover la redundancia temporal y codificación por transformación de la señal resultante para reducir la redundancia espacial es adoptada. El decodificador tiene la capacidad de utilizar la técnica de compensación de movimiento, permitiendo la incorporación opcional de esta técnica en el codificador.

Cada bloque de 8×8 en una imagen es transformado en coeficientes de transformada de coseno discreto, cuantizados linealmente, y los datos resultantes son codificados mediante codificación de longitud variable y los datos agrupados en la estructura de datos jerárquicos referida. Esta estructura de datos es enviada al almacenador de transmisión (del inglés *buffer transmission*), el cual regula el flujo de información de video a una tasa de bit constante al controlar el nivel de cuantización del cuantizador lineal. El nivel de cuantización puede ser ajustado dependiendo del contenido del almacenador de transmisión del codificador. Cuando el almacenador de transmisión está cerca de llenarse, el valor de cuantización será incrementado de tal manera que se produzca menos información binaria en el proceso de codificación. Esto resulta en una imagen degradada. Por otra parte, el nivel de cuantización será decrementado para mejorar la calidad de la imagen cuando el almacenador de transmisión no esté lleno.