



Fundamentos de Voz y Video sobre IP

Preparado por: Jorge Olivares A.

2014

Índice



Voz y video sobre IP (parte 1)	2	Fotogramas P (Predictivos)	58
Naturaleza de los sonidos y la voz	3	Fotogramas B (Bidireccionales)	59
Transductores de audio	6	Caudal de una videoconferencia	60
Generación de la voz	7	Audio MPEG-1	61
Fisiología del sistema auditivo	8	Sincronización de audio y vídeo MPEG	62
Modelo psicoacústico de la Audición	9	Vídeo MPEG-2 (1)	63
Umbrales de la Audición	15	Vídeo MPEG-2 (2)	64
Formatos de audio	16	Caudales de Niveles y Perfiles MPEG-2	65
Luz, color, imagen	17	Formatos de imagen	66
Los principios de la percepción de las imágenes	19	Formatos de video digital	67
El ojo y los principios de la percepción de las imágenes	20	Multiplexación	68
La luz y los colores	22	Multiplex MPEG	69
Generación de la imagen digital (cámara)	23	Formación del paquete MPEG (PS)	70
Tecnologías de cámaras digitales (CCD y CMOS)	24	Formación de la trama MPEG-TS	71
Vídeo digital	26	Voz y video sobre IP (parte 3)	72
La percepción del movimiento	27	Clasificación del Servicio Telefónico	73
Pixel y resolución de imágenes	28	La naturaleza de las redes telefónica y de datos	74
Resolución en impreso y pantalla	29	Convergencia de las redes telefónica y de datos	75
Mapa de bits versus vectorial	30	Conmutación de circuitos	76
Resolución de pantalla	31	Conmutación de paquetes	77
Formatos de pantalla en televisión	33	Cx Circuitos versus Cx de paquetes	79
HDTV (Alta Definición)	34	Modelo OSI de una red de datos	80
La formación de la imagen en TV y computación	35	Protocolos: TCP	81
Voz y video sobre IP (parte 2)	36	Protocolos: UDP	82
Procesamiento de la información en un sistema de telecomunicaciones	37	Calidad de Servicio (QoS)	84
Fundamentos de transmisión	38	Parámetros de Calidad de Servicio	85
Procesamiento de la señal para la transmisión	39	Efectos de la congestión	86
Codificación de la Voz	40	Relación entre la probabilidad de llegada de los paquetes y los parámetros de QoS	87
Conversión análogo-digital de la voz	41	Concepto de Jitter	88
Codificación de la Voz	42	Reducción del Jitter	89
Generación de bits de imagen de video y compresión	43	Requerimientos de Calidad de Servicio de las aplicaciones	90
Compresión de video	44	¿Reserva o Prioridad?	91
Compresión de video sin pérdidas	45	Calidad de Servicio en Internet	92
Submuestreo (compresión de vídeo con pérdidas)	46	Protocolo RTP	93
Submuestreo 4:2:2	47	VoIP - ToIP	95
Submuestreo 4:1:1	48	Voz y video sobre IP (parte 4)	96
Submuestreo 4:2:0	49	Estandarización de la VoIP	97
Formatos compresión de vídeo	50	Los estándares de video y televisión analógica	98
Caudal requerido por los sistemas de compresión de vídeo más comunes	51	Estándar NTSC	99
Vídeo M-JPEG (Motion JPEG)	52	Resolución de video para NTSC	100
Funcionamiento de MPEG	53	Evolución de estándares de video	101
Vídeo MPEG (MPEG-1)	54	Formatos y estándares de video digital	102
Vídeo MPEG (tipos de fotogramas)	55	Vídeo H.26x	103
Fotogramas I (Intra)	56	Estándar H264 de ITU	104
	57	Aspectos regulatorios de telefonía IP en Chile	106

1: Principios físicos de imagen y sonido y fisiológicos de audición y visión humana y su conversión para el procesamiento y transmisión

2: Procesamiento y transmisión analógico y digital de señales de voz, audio, imagen y video

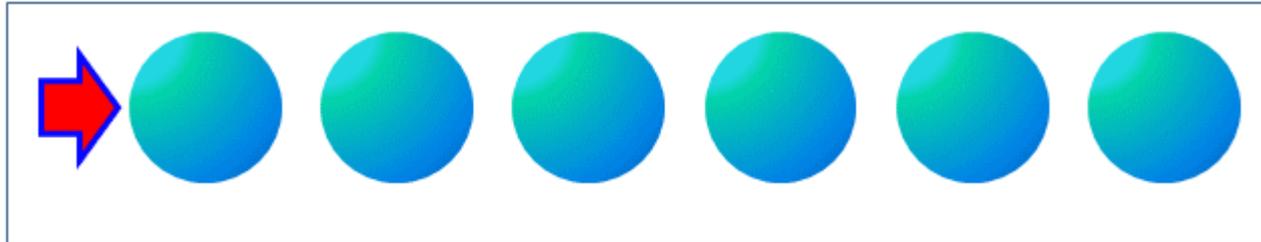
3: Características y requerimientos de una red IP para el transporte de señales de voz, audio y video

4: Estándares y normativa de sistemas de procesamiento y transmisión de voz y video IP

Naturaleza de los sonidos y la voz



Concepto: Sonido es una vibración molecular de un medio elástico (por ejemplo el aire) que se propaga en forma de una onda mecánica de presión y que se rige por las leyes de la Física Acústica.



El sonido se propaga por el aire a una velocidad de 340 m/seg. y se atenúa con la distancia a medida que se propaga en el espacio.

En la atmósfera terrestre se propaga como fluctuaciones de la presión atmosférica, tomando valores por encima y por debajo del valor estático. Se define Presión acústica (sonora) a:

$$p = P - P_0$$

Donde:

P: Presión atmosférica total medida en N/m^2 (Newton por metro cuadrado)

P_0 : Presión atmosférica estática medida en N/m^2 ($= 1,013 \cdot 10^5 \text{ N/m}^2$ al nivel del mar y a 0°C)

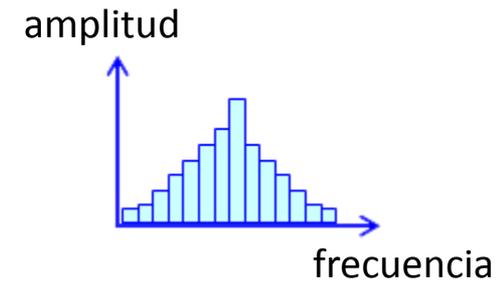
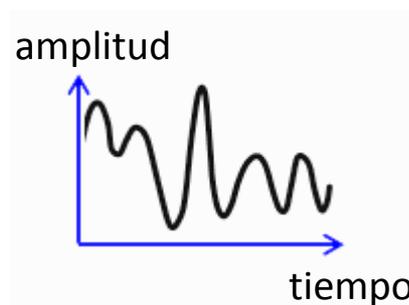


Naturaleza de los sonidos y la voz



Una onda sonora es una variación de presión acústica del aire. La onda sonora esta formada por múltiples componentes (Fourier) que se pueden observar en un osciloscopio (su forma de onda) y en un analizador de espectro (componentes de frecuencia)

$$p = \sum_{n=1}^N A_n \text{ sen } (2\pi n f_0 t)$$



Cada componente se denomina tono, ya que involucra una sola frecuencia.

Actividad práctica: Para quienes disponen de dispositivos con sistema operativo Android, bajar de la tienda la aplicación Spectrum Analyzer o equivalente. Hablar o emitir sonidos. Observar que se visualiza cuando se emite un silbido constante.

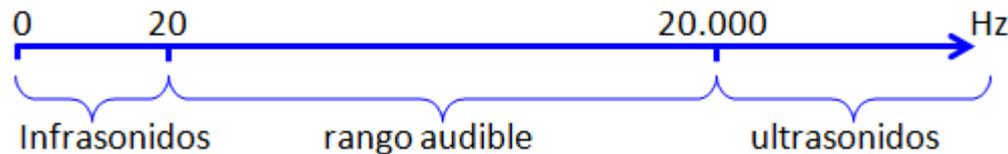


Naturaleza de los sonidos y la voz



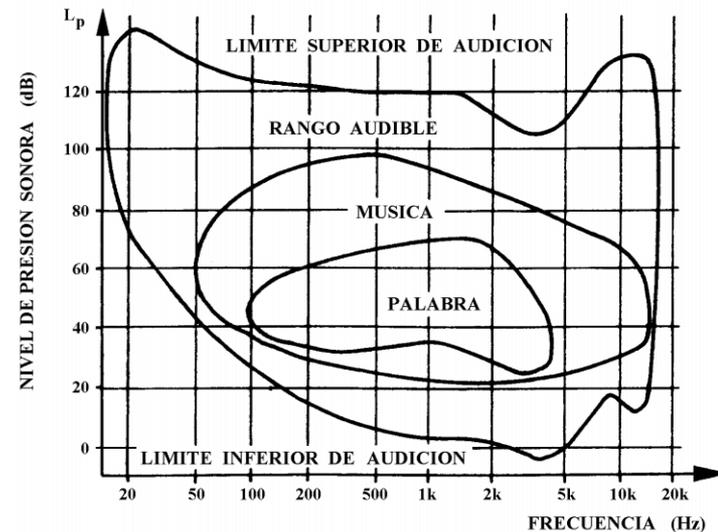
El sonido puede ser audible o no para el ser humano. De hecho, esta condición es variable entre los individuos y en particular a lo largo de su vida, disminuyendo la capacidad en forma progresiva con la edad.

Sonidos no audibles para el ser humano se localizan bajo frecuencias de 20 Hz (infrasonidos) y sobre 20.000 Hz (ultrasonidos)



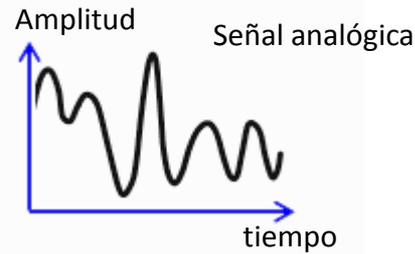
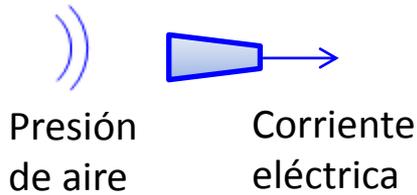
Dentro del rango audible se distingue la voz de canal telefónico (BW = 4 KHz), el audio de recepción de radiodifusión AM (BW = 7 KHz y FM (BW = 15 KHz)

Actividad práctica: Para quienes disponen de dispositivos con sistema operativo Android, bajar de la tienda la aplicación Generador de tonos FG o equivalente. Realizar audición a diferentes frecuencias.



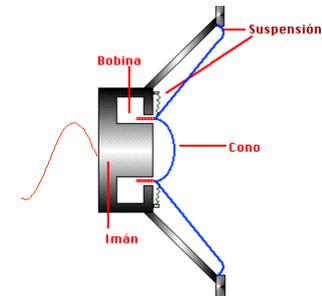
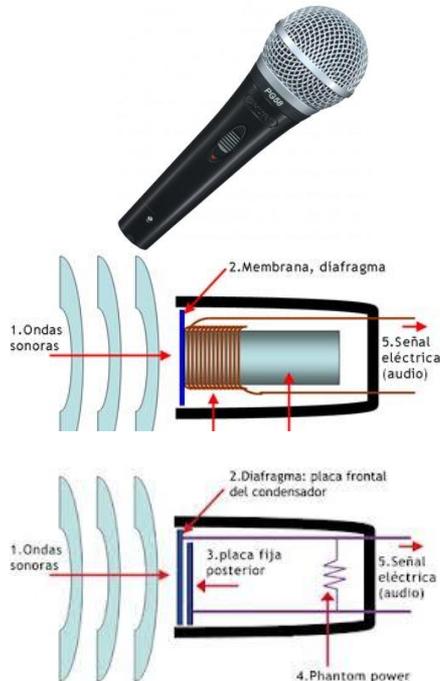
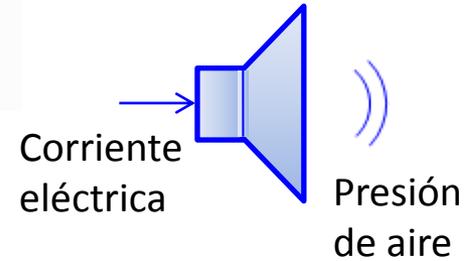
Transductores de audio

micrófono:



Voz
Sonidos
Música
Ruido

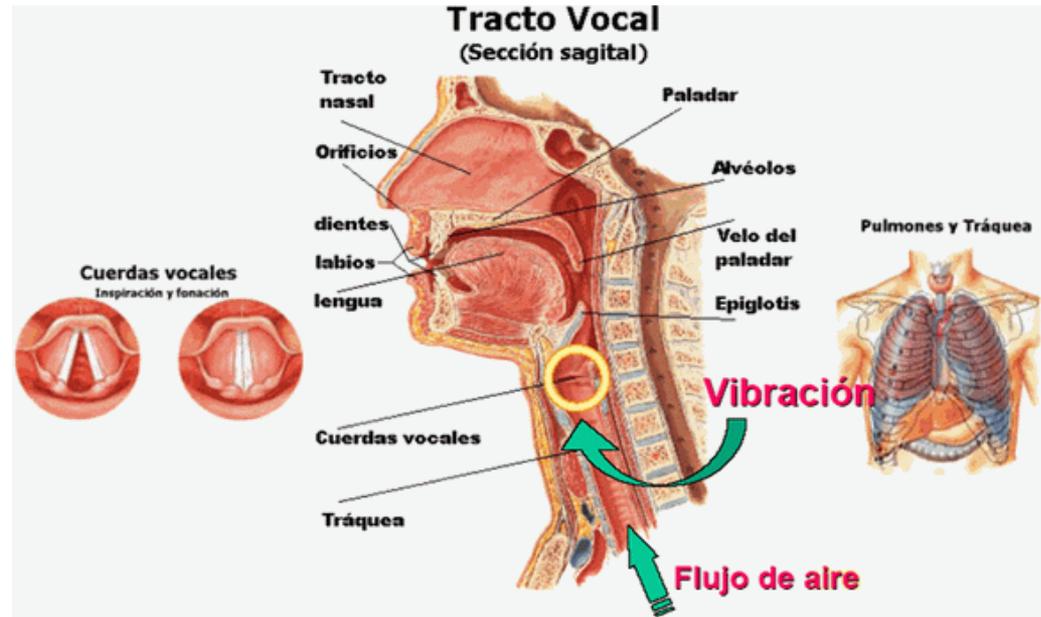
parlante:



Generación de la voz

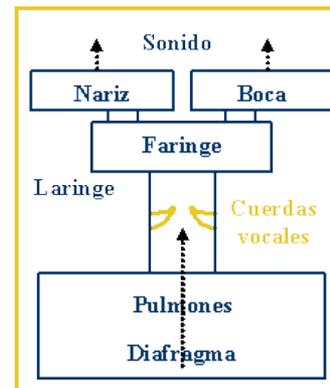


Concepto: La voz se genera como un sonido proveniente de complejo proceso donde el flujo de aire que se expulsa de los pulmones pasa por una membranas denominadas cuerdas vocales que vibran y transfieren esta vibración al flujo de aire que se propaga desde la boca.



El contenido de información va codificado mediante el lenguaje que los individuos aprenden y transfieren de generación en generación.

El sistema vocal también puede producir sonidos especiales como por ejemplo el grito, el llanto, silbido, etc.



El sistema vocal humano puede dividirse en tres partes:

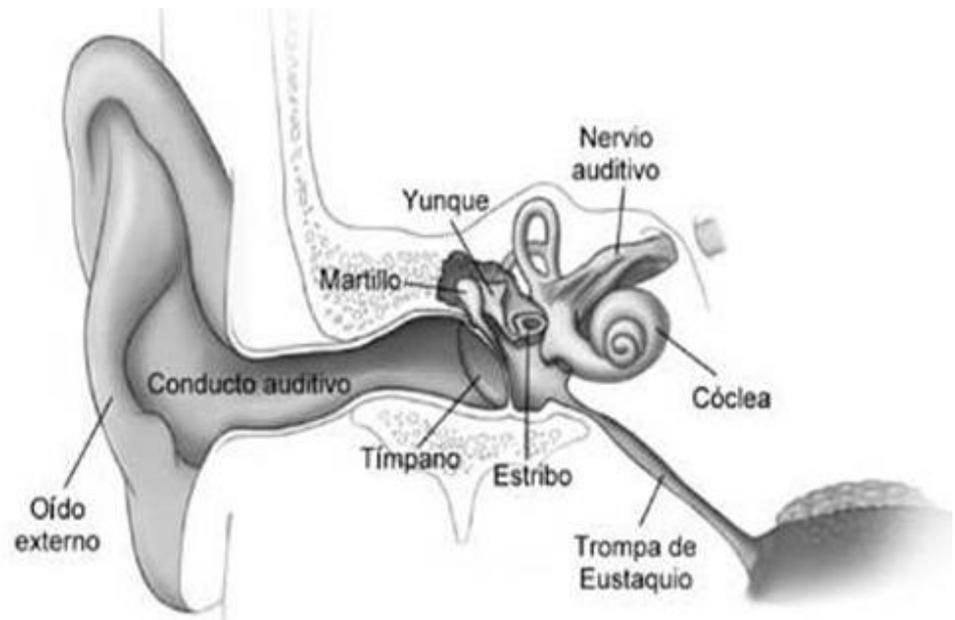
- **Aparato respiratorio:** donde se almacena y circula el aire. Nariz, traquea, pulmones y diafragma.
- **Aparato de fonación:** donde el aire se convierte en sonido. Laringe y cuerdas vocales.
- **Aparato resonador:** donde el sonido adquiere sus cualidades de timbre que caracterizan cada voz. Cavidad bucal, faringe, paladar óseo, senos maxilares y frontales.

Fisiología del sistema auditivo



Las ondas sonoras ingresan al oído externo y viajan a través del conducto auditivo hasta el tímpano.

El tímpano vibra con las ondas sonoras y transmite estas vibraciones a lo largo de tres huesos denominados martillo, yunque y estribo que se ubican en el oído medio. Estos amplifican el sonido y lo envían a través de la entrada del oído interno (ventana oval) hacia el órgano auditivo (cóclea), cubierto por fluido.



Las vibraciones crean ondas en el líquido que estimulan las terminaciones nerviosas de la cóclea (células ciliadas) causando impulsos eléctricos que el nervio auditivo envía al cerebro.

El cerebro traduce estos impulsos en lo que percibimos como sonido.

Modelo psicoacústico de la Audición



Concepto: El contenido de información en los sonidos va codificado mediante el lenguaje o sentido asignado a sonidos como por ejemplo la música u otros cuyo significado se asigna en forma arbitraria (por ejemplo una sirena, un ringer, una campana, etc.)

Por otra parte, el complejo proceso de la audición muestra que existen condiciones en las cuales no todos los sonidos tienen la misma relevancia.

Estas propiedades son consideradas en los sistemas de audio y en particular por los de compresión para disminuir la cantidad de datos necesarios para representar un sonido, basándose en un modelo psicoacústico que simula el comportamiento del oído humano.

Se toma ventaja de las limitaciones del oído (así como del cerebro) para responder a todas las componentes en una onda de audio compleja y, de esta manera, lograr que los mecanismos de compresión determinen lo que se oirá de un sonido particular, descartando los sonidos indetectable o codificándolo con menos precisión, de manera de reducir la cantidad de información a procesar y transportar sin cambiar la calidad del sonido percibido.

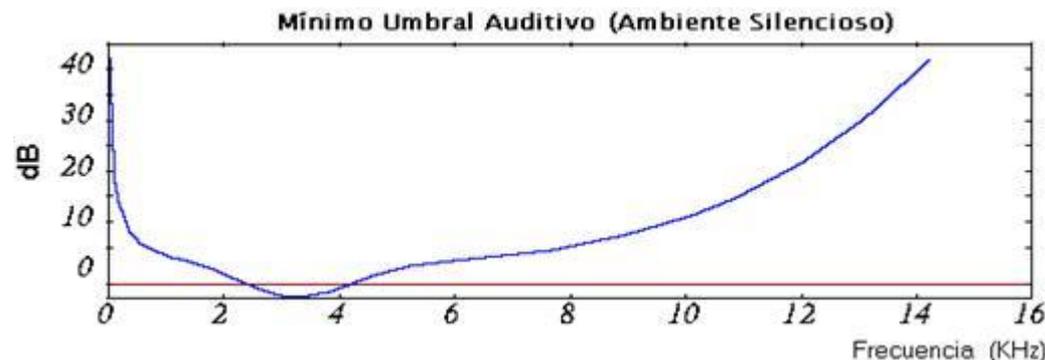
Modelo psicoacústico de la Audición



Existen cuatro fenómenos que son los más importantes y útiles para el desarrollo de los codificadores de audio.

1. Mínimo umbral auditivo

Este umbral, también conocido como umbral absoluto, corresponde al sonido de intensidad más débil que se puede escuchar en un ambiente silencioso. El mínimo umbral auditivo no tiene un comportamiento lineal; se representa por una curva de Intensidad (dB) contra Frecuencia (Hz), que posee niveles mínimos entre 2 y 5 KHz, los cuales corresponden a la parte más sensitiva del oído humano. Por lo tanto, en los sistemas de compresión de audio que sacan provecho de la psicoacústica, no es necesario codificar los sonidos situados bajo este umbral (el área por debajo de la curva), ya que éstos no serán percibidos.



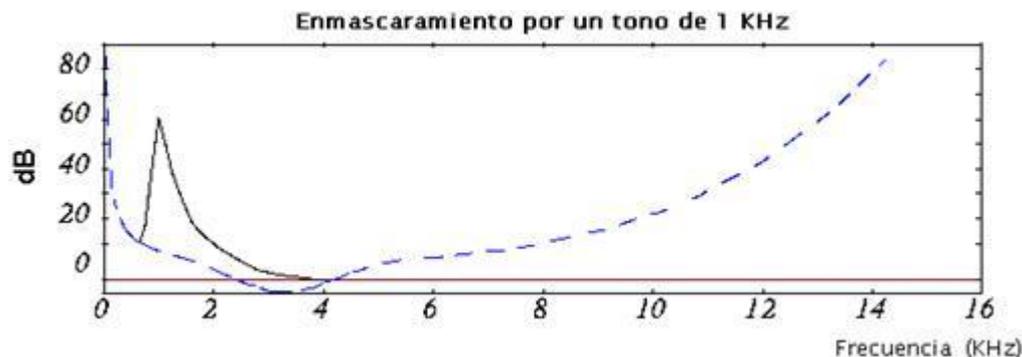
Fuente: <http://www.labc.usb.ve/paginas/EC4514/AUDIO/PSICOACUSTICA/Psicoacustica.html>

2. Enmascaramiento.

El efecto de enmascaramiento se basa en las limitantes del oído humano para responder a todas las componentes de un sonido complejo. Durante los sonidos fuertes, no se pueden oír los sonidos más débiles. Por ejemplo, cuando un músico organista no está tocando, se puede escuchar el resoplido de los tubos; y cuando el músico toca, se pierde el sonido de éstos porque ha sido enmascarado. Es el ejemplo más fácil de entender de enmascaramiento.

2.1. Enmascaramiento en frecuencia.

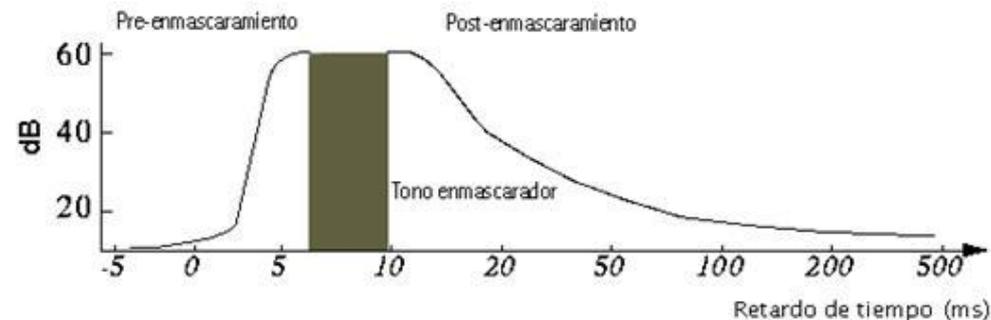
Funciona de manera que un sonido en determinada frecuencia puede enmascarar o disminuir el nivel de otro sonido en las frecuencias adyacentes, siempre y cuando el nivel del sonido enmascarante sea más alto (un sonido más intenso, más fuerte) que el nivel del sonido adyacente.



Fuente: <http://www.labc.usb.ve/paginas/EC4514/AUDIO/PSICOACUSTICA/Psicoacustica.html>

2.2. Enmascaramiento temporal.

Se presenta cuando un tono suave está muy cercano en el dominio del tiempo (unos cuantos milisegundos) a un tono fuerte. Si se está escuchando un tono suave y aparece un tono fuerte, el tono suave será enmascarado por el tono fuerte, antes de que el tono fuerte efectivamente aparezca (preenmascaramiento) en alrededor de 30 ms antes. Posteriormente, cuando el tono fuerte desaparece, el oído necesita un pequeño intervalo de tiempo (entre 50 y 300 ms) para que se pueda seguir escuchando el tono suave (post-enmascaramiento).



El preenmascaramiento sugiere que un tono será enmascarado por otro tono, antes de que el tono enmascarador realmente aparezca, atentando contra el buen juicio de cualquier oyente. Para este fenómeno, se han presentado dos explicaciones:

- 1) El cerebro integra el sonido sobre un período de tiempo, y procesa la información por ráfagas en la corteza auditiva
- 2) El cerebro procesa los sonidos fuertes más rápido que los sonidos suaves.



3. *Joint stereo* o estéreo conjunto

Si una señal de audio es estereofónica se puede lograr comprimirla, con base en la irrelevancia o redundancia entre ambos canales. Por ejemplo, en muchas locaciones de alta fidelidad, existe un único y potente altavoz denominado "*BOOMER*".

Sin embargo, se tiene la impresión de que el sonido proviene de diferentes fuentes como si existieran parlantes en todas las direcciones. En la realidad, por debajo de una frecuencia determinada, el oído ya no es capaz de localizar el origen espacial de los sonidos.

De esta manera, algunas frecuencias se pueden grabar como señal monofónica seguida por un pequeño código para lograr restaurar un pequeño porcentaje de espacialización en la decodificación.



4. Bandas críticas.

Estudios de la discriminación en frecuencia del oído demuestran que en bajas frecuencias, tonos con unos cuantos Hertz de separación pueden ser distinguidos; sin embargo, en las altas frecuencias para poder discriminar los tonos se necesita que estén separados por cientos de Hertz. En cualquier caso, el oído responde al estímulo más fuerte que se presente en sus diferentes regiones de frecuencia; a este comportamiento se le da el nombre de bandas críticas.

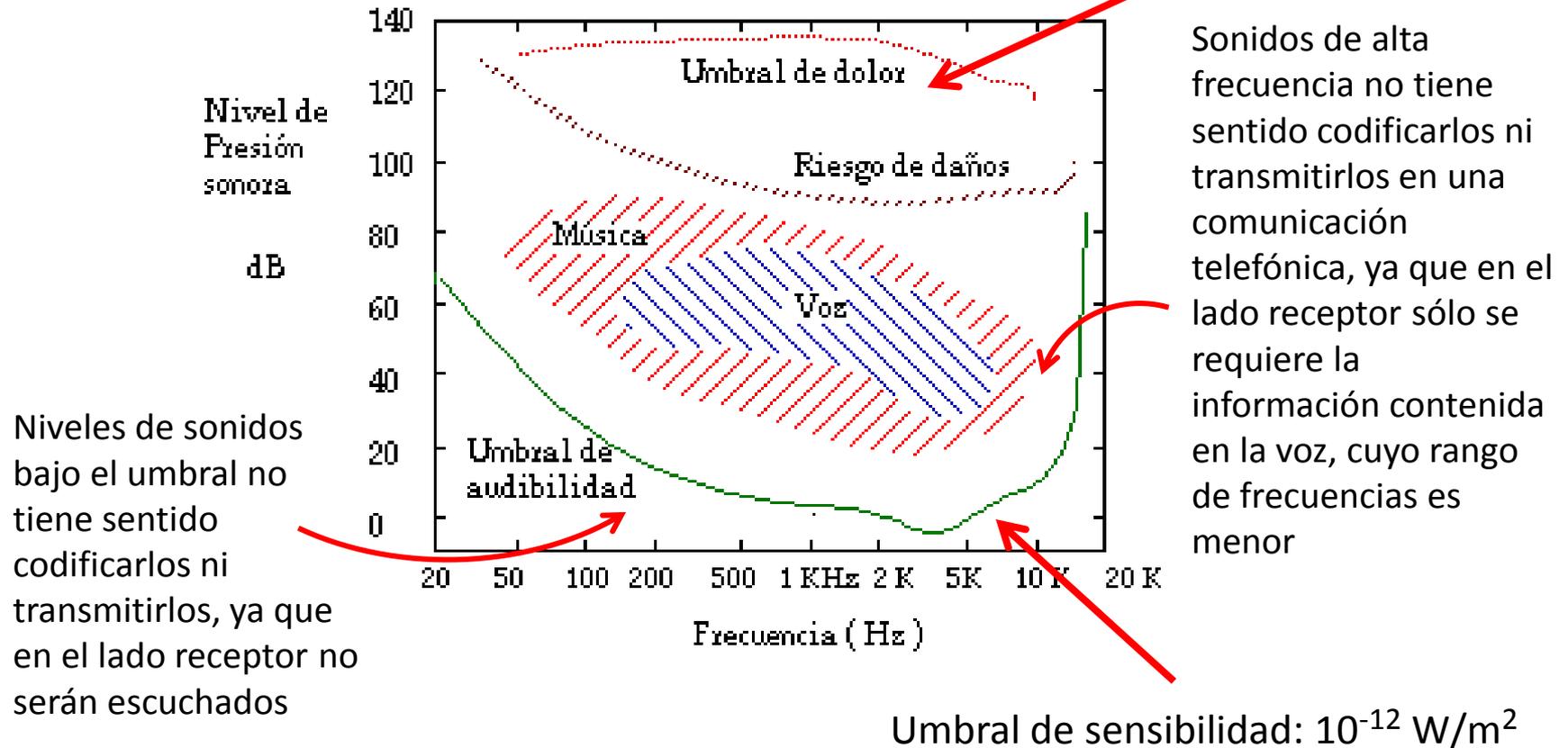
Los estudios muestran que las bandas críticas son más estrechas en las bajas frecuencias que en las altas; el 75% de las bandas críticas están por debajo de los 5 KHz, lo que implica que el oído recibe más información en las bajas que en las altas frecuencias. Las bandas críticas tienen un ancho de aproximadamente 100 Hz para las frecuencias de 20 a 400 Hz; este ancho aumenta de manera logarítmica a medida que aumenta la frecuencia.

Más importante es el hecho de que las bandas críticas no son fijas; son continuamente variables en frecuencia y cualquier tono audible creará una banda crítica centrada en él. Mirado desde otro punto de vista, el concepto de la banda crítica es un fenómeno empírico: una banda crítica es el ancho de banda al cual las respuestas subjetivas cambian abruptamente.

Umbrales de la Audición

Área de audición y umbrales mínimos y máximos de audición:

Dolor: $1,0 \text{ W/m}^2$ (120 dB)



Nivel de referencia: 10^{-12} W/m^2 (0 dB)

Formatos de audio



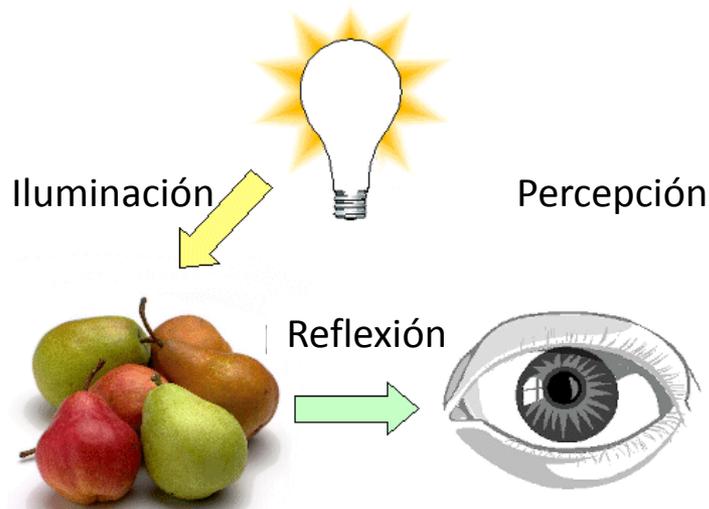
Formato	Descripción	Extensión
Waveform Audio File Format (WAV)	Formato estándar sin compresión para audio sin pérdidas. Tamaño grande de archivo.	.wav
MPEG-1 o MPEG-2 Audio Layer 3 o III (MP3)	Popular formato con compresión y pérdida de calidad de audio. Elimina las partes menos “humanamente” audibles. Archivos pequeños.	.mp3
Windows Media Audio (WMA)	Por defecto para Windows Media Player. Utiliza algoritmos propietarios para comprimir audio.	.wma
Xiph.org Open Standard (OGG)	Formato abierto que soporta una variedad de codecs de audio, siendo el más popular Vorbis. Compresión similar a mp3.	.ogg, .oga, .ogx, .ogg, .spx
Advanced Audio Coding (AAC)	Basado en MPEG-2 y MPEG-4. Una versión que incluye derechos de autor es propiedad de apple y usada en su tienda iTunes Music Store. La extensión de esta variación es .m4p	.m4a, .m4b, .m4p, .m4v, .m4r, .3gp, .mp4, .aac

Luz, color, imagen



Concepto: La visión es uno de nuestro sentidos que nos permite observar lo que nos rodea mediante la luz.

La Física aún discute acerca de la naturaleza de la luz, ya que, en ciertas situaciones presenta comportamiento corpuscular (partícula) y en otras presenta comportamiento ondulatorio (ondas)



La luz del sol y de fuentes artificiales de luz generalmente es blanca y en parte es absorbida o reflejada por la superficie de los objetos.

La luz blanca esta compuesta por una amplia gama de haces de diferentes longitudes de onda (colores).

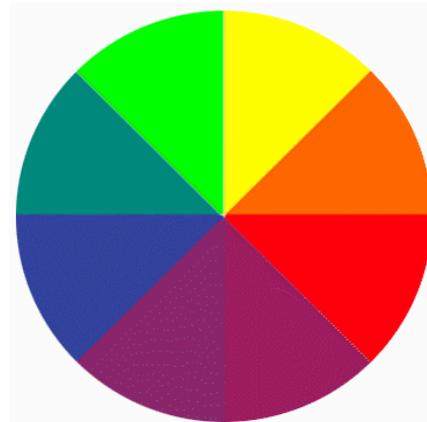
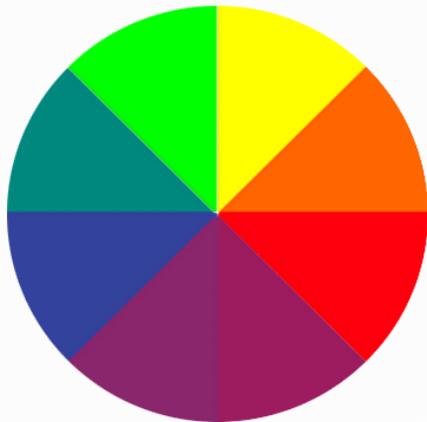
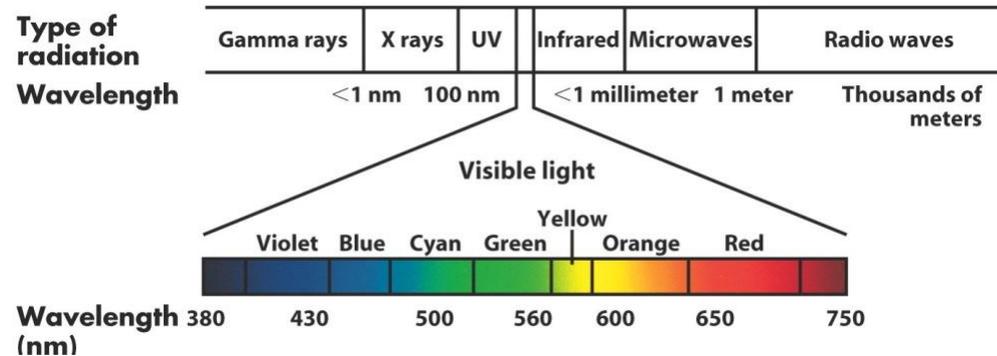
Un objeto de “color verde” absorbe todos los colores que forman la luz blanca, excepto el color verde qué lo refleja y permite que esa luz llegue a nuestro ojo.

Luz, color, imagen



Luz blanca esta compuesta por múltiples colores (diferentes longitudes de onda).

El espectro visible se encuentra localizado entre los 380 nm (nanómetro) y los 750 nm.
 $1\text{nm} = 1 \times 10^{-9} \text{ m}$.



Disco de Newton permite demostrar que la luz blanca se obtiene a partir de la combinación de los colores básicos.

Los principios de la percepción de las imágenes



Lo que vemos esta directamente relacionado con la percepción del color de la luz que reflejan los objetos o que producen las fuentes luminosas.

Los principales parámetros que definen el color percibido en una imagen son:

- ✓ **Tono (hue):** distingue entre las diferentes tonalidades (rojo, azul, amarillo, etc.). Este parámetro esta asociado directamente a la longitud de onda.
- ✓ **Saturación:** medida de la pureza del color; los colores no saturados contienen una proporción de luz blanca. Los colores saturados no contienen luz blanca.
- ✓ **Luminosidad (lightness):** intensidad percibida desde un objeto que refleja la luz. Este parámetro esta relacionado con la intensidad o nivel de potencia.
- ✓ **Brillo (brightness):** similar a la luminosidad, pero referido a objetos que emiten luz (sol, lámpara, pantallas).

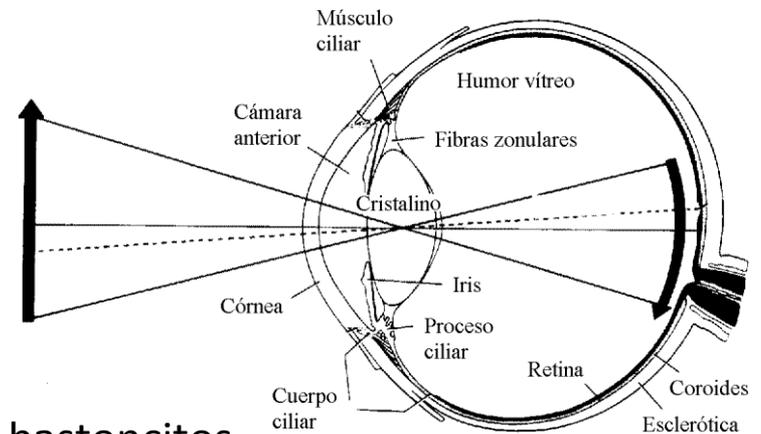
Otros aspectos relacionados son: resolución de imagen, relación de aspecto, cantidad de colores, etc.

El ojo y los principios de la percepción de las imágenes



Cuando el ojo está bien enfocado, la luz reflejada por un objeto exterior forma su imagen en la retina.

El mecanismo o patrón de la visión está basado en la distribución de receptores de luz discretos sobre la superficie de la retina.



Se encuentran dos tipos de receptores: conos y bastoncitos.

Los conos totalizan entre 6 y 7 millones, con una fuerte concentración en la región central de la retina llamada fovea, siendo altamente sensitivos al color. La capacidad de resolución de los detalles finos de una imagen se debe fundamentalmente a que cada uno de estos conos cuenta con su propia terminación nerviosa.

Los conos se distribuyen en 64% sensibles a la luz roja, 32% sensibles al verde y solo el 2% sensible a la luz azul.

Los conos sensibles al azul tienen mayor sensibilidad a la luz, que los sensibles a la luz roja y verde, pero hay menos de ellos.

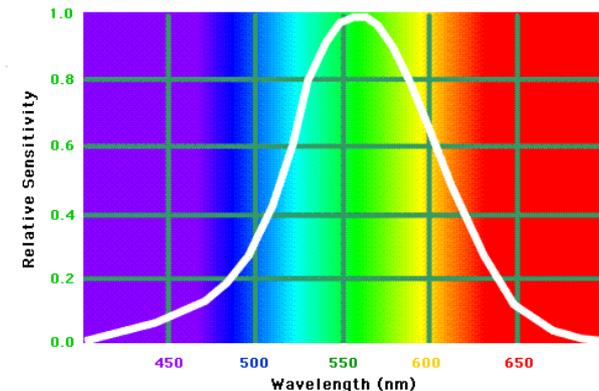
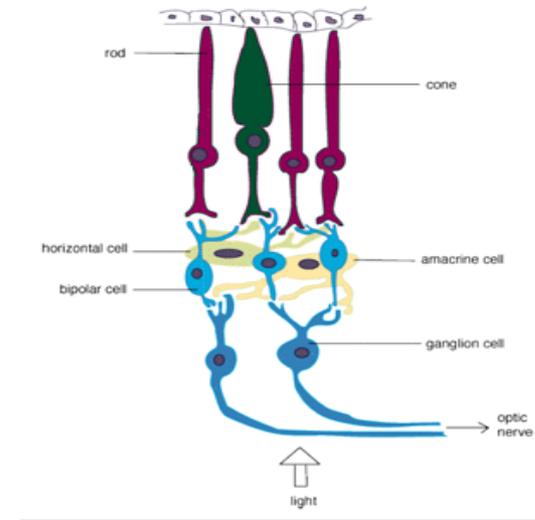
El ojo y los principios de la percepción de las imágenes



Si observamos tres objetos que irradian la misma cantidad de luz, el ojo humano percibirá el objeto verde como el más brillante, después el rojo y el azul como el menos brillante. Esto queda representado por la curva de sensibilidad relativa que muestra su nivel más alto entre el verde y amarillo.

Los bastoncitos son más de 100 millones y se distribuyen en un área mayor y el hecho de estar varios bastoncitos conectados a terminaciones nerviosas comunes, reducen el grado de detalle que es posible discernir. Por esto podemos concluir que:

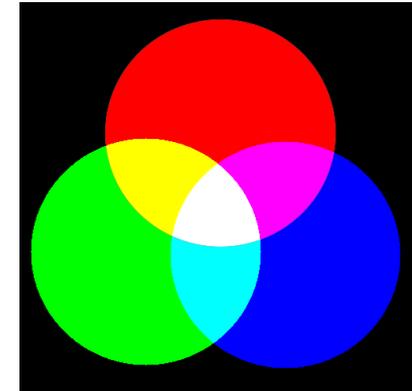
- ✓ Los bastoncitos no son sensibles al color, pero si a niveles bajos de intensidad de luz.
- ✓ Los conos son sensibles al color y a niveles más altos de intensidad de luz.



Fuente: http://people.deas.harvard.edu/~jones/cscie129/nu_lectures/lecture8/color/luminosity.html

La luz y los colores

Los colores primarios en la luz son el rojo (R), verde (G) y azul (B). El resto de los colores se obtiene por la combinación de los anteriores. Este principio es el que gobierna todos los sistemas de generación y visualización de imágenes (cámaras y monitores).



Para obtener el color blanco es necesario efectuar una combinación de los colores básicos en la proporción que determina la sensibilidad de la visión humana a los diferentes colores.

En particular, la siguiente combinación de colores se percibe como luz blanca y se denomina luminancia : $Y = 0,3 \times R + 0,59 \times G + 0,11 \times B$

Por otra parte y dependiendo del nivel de potencia de la luminancia Y se puede obtener la escala de grises.

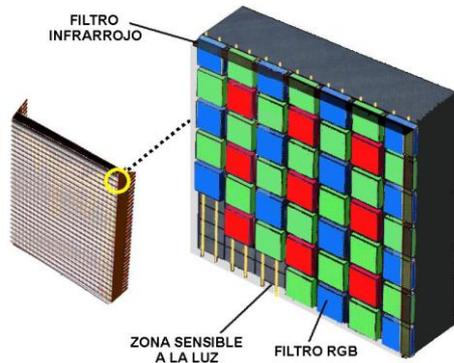
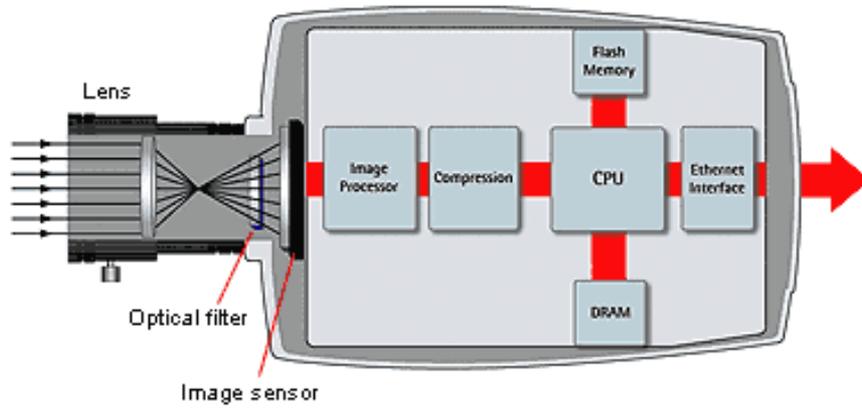


Cualquier otra combinación de los colores R, G y B genera un determinado color. Esto se denomina Crominancia. A continuación se muestran combinaciones donde cada color tiene un rango de 0 a 255 niveles de saturación.

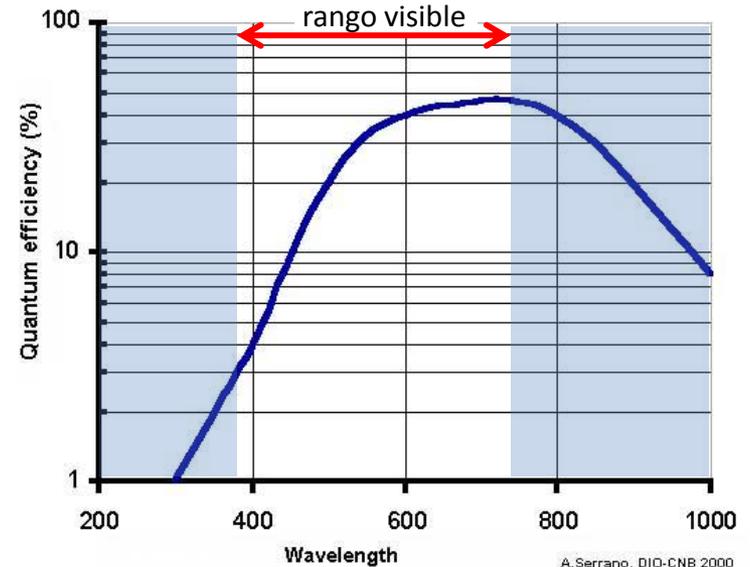
R= 255	R= 253	R= 209	R= 8	R= 2
G= 255	G= 250	G= 3	G= 249	G= 2
B= 255	B= 3	B= 253	B= 218	B= 2

Generación de la imagen digital (cámara)

Diagrama general de una cámara de video



Para obtener los colores RGB se recurre a microscópicos filtros



Curva de sensibilidad de un CCD
Se observa que el rango se extiende más allá del espectro visible (380 a 750 nanómetros). Esto explica la existencia de las cámaras de visión nocturna o infrarrojas.

Actividad práctica: Mediante la cámara de su celular observar la luz que emite el led del control remoto (mando) de cualquier dispositivo del hogar.

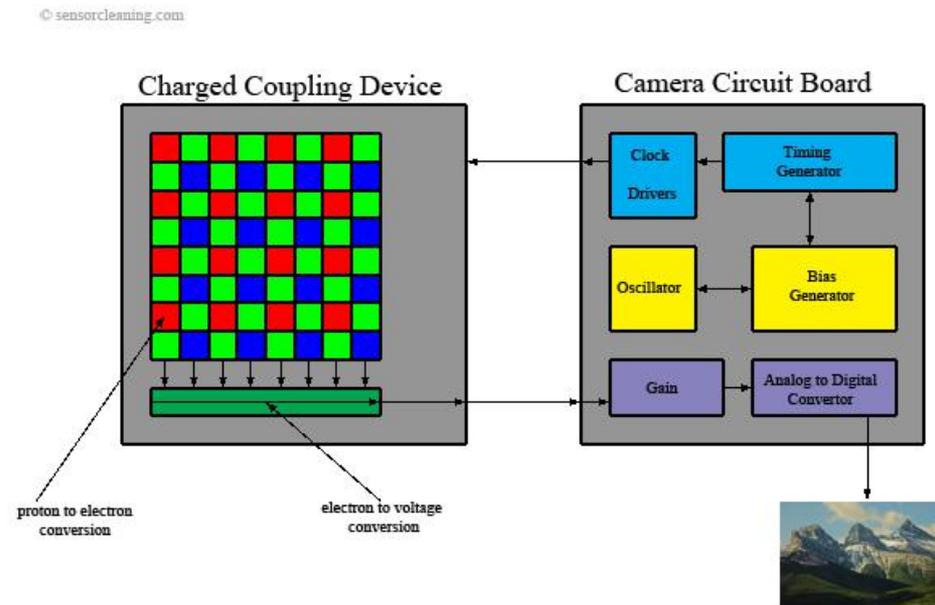
Tecnologías de cámaras digitales (CCD y CMOS)



En la actualidad la mayoría de las cámaras fotográficas digitales y de video se fabrican sobre la base de tecnología CCD y CMOS.

CCD: del inglés Charge Coupled Device (dispositivo de carga acoplada).

- Este tipo de sensor transmite la información de cada célula a las contiguas en filas y columnas hasta los laterales del sensor, donde se recoge esta información y se envía al conversor analógico-digital fuera del propio sensor, donde dicha información es convertida a datos aptos para ser guardados en la tarjeta de memoria.
- Sus ventajas son un menor ruido de salida, una mayor sensibilidad por pixel y los años que lleva siendo usado y mejorado en elementos de imagen digital.
- Sus desventajas son su mayor consumo eléctrico (con un aumento de temperatura y necesidad de refrigeración) y su menor rapidez de captación comparado con el CMOS.

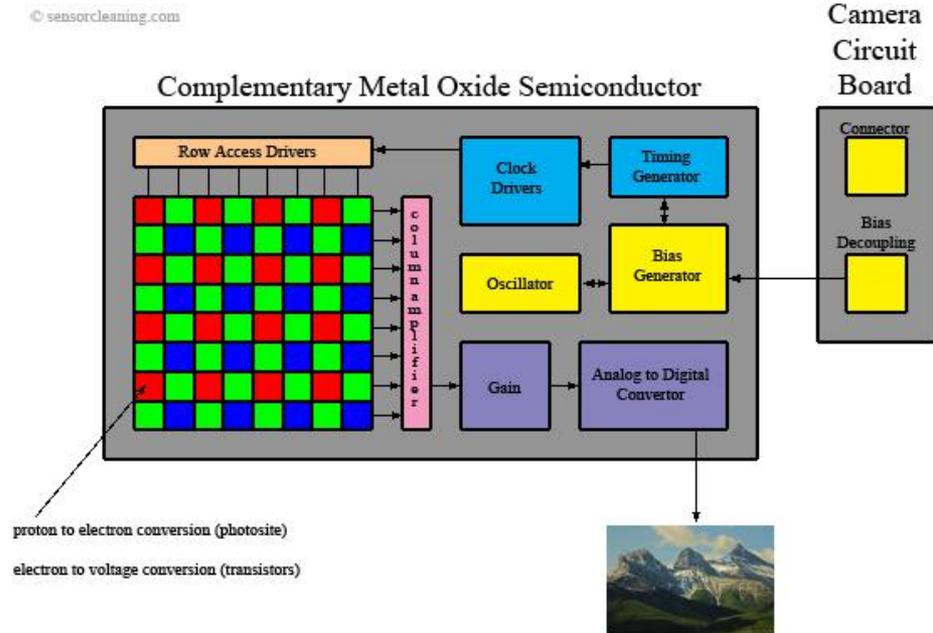


Fuente: <http://www.fotov60.com/2011/02/12/ccd-y-cmos>

Tecnologías de cámaras digitales (CCD y CMOS)



- CMOS: Del inglés complementary metal oxide semiconductor, “semiconductor de metal oxido complementario”.
- Este tipo de sensor incorpora un conversor analógico-digital incorporado en cada pixel, acelerando la conversión de la información y su conversión a datos aptos para ser guardados en la tarjeta de memoria.
- Sus ventajas son la posibilidad de capturar imagen a mayor velocidad gracias a su estructura, una menor temperatura de trabajo y un menor costo de fabricación.
- Su desventaja es un mayor ruido de salida y una menor sensibilidad debido a que entre los receptores de luz (fotositos) se encuentra mucha electrónica que no es sensible a la luz.



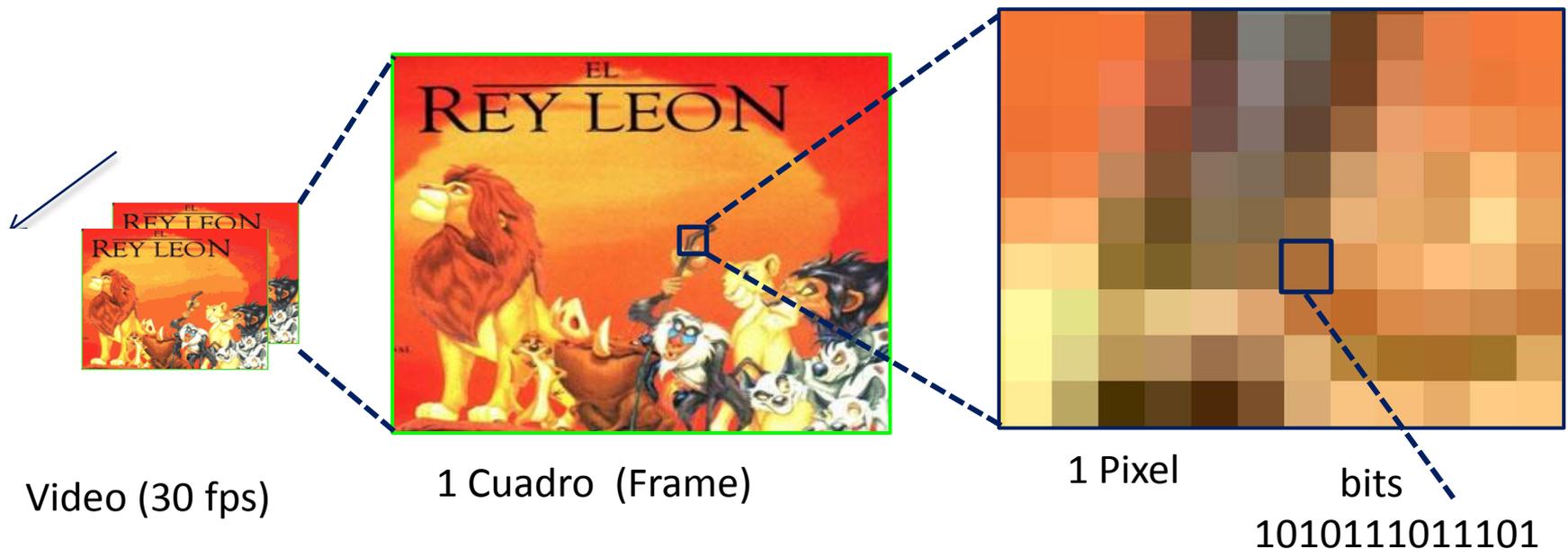
Actualmente las diferencias entre ambas tecnologías se han reducido notablemente, debido al interés de los fabricantes en los CMOS por su menor costo de fabricación y consumo eléctrico. La tendencia actual es mejorar dicho sensor e incorporarlo al máximo número de cámaras.

Fuente: <http://www.fotov60.com/2011/02/12/ccd-y-cmos>

Video digital



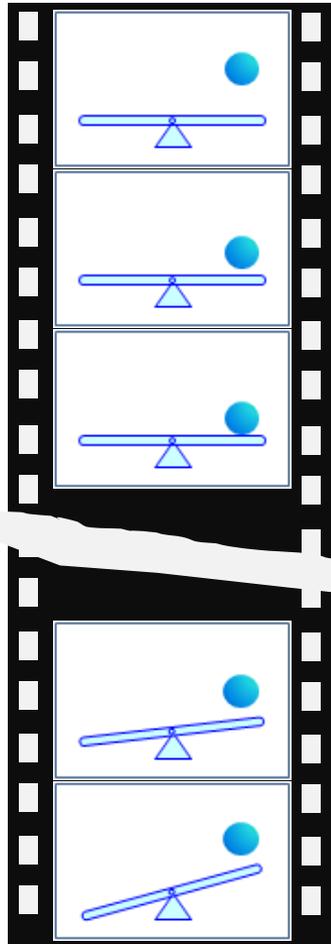
- Una señal de video esta compuesta por 30 cuadros de imagen (30 fps)
- Un cuadro de imagen esta compuesto por una cierta cantidad de pixeles
- (STD 720x480, HDTV 1920x1080)
- Cada pixel se asocia a una cierta cantidad de bits que representan un nivel de brillo y color (8, 10, 16, 24 bits) y el video queda definido por un tren de bits continuos.



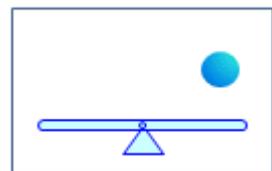
La percepción del movimiento



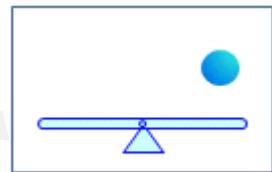
La imagen en movimiento es una ilusión óptica, ya que corresponde a la presentación de una secuencia de imágenes fijas mostradas a cierta velocidad en que la visión humana es incapaz de separarlas y el cerebro las integra en una sola.



En general a 25 cuadros por segundo (fps : frames per seconds) la visión humana tiene dificultades para discriminar entre imágenes fijas, por lo cuál se ha adoptado 30 fps como el estándar que asegura la percepción plena del movimiento y eliminación del parpadeo por cambio de imágenes entre cuadros.



12,5 fps.



4,2 fps



1,8 fps

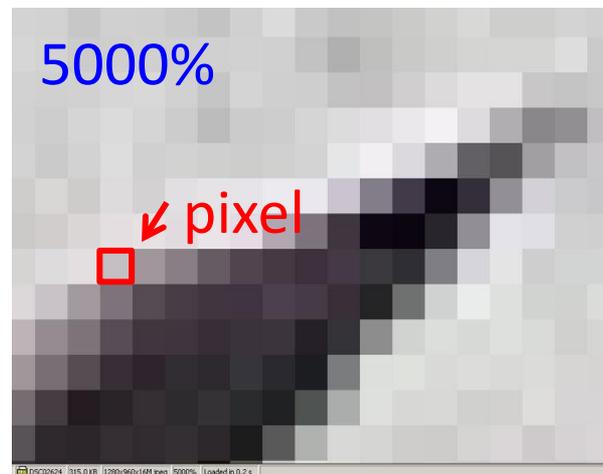
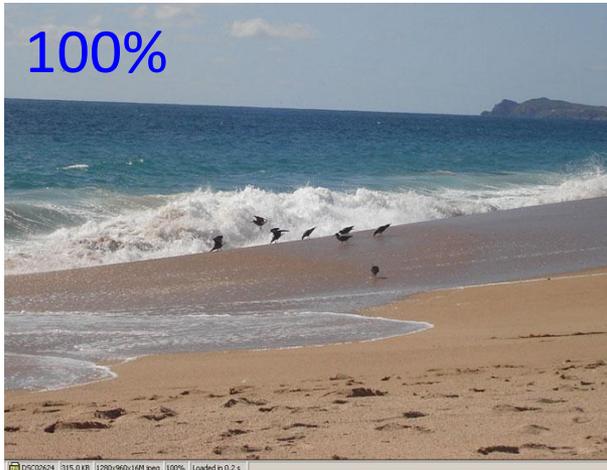
Más fps implica mayor volumen de información y por lo tanto mayor ancho de banda o velocidad para procesar las imágenes en tiempo real

Este principio rige para el cine, televisión, computación, videoconferencia y videovigilancia.

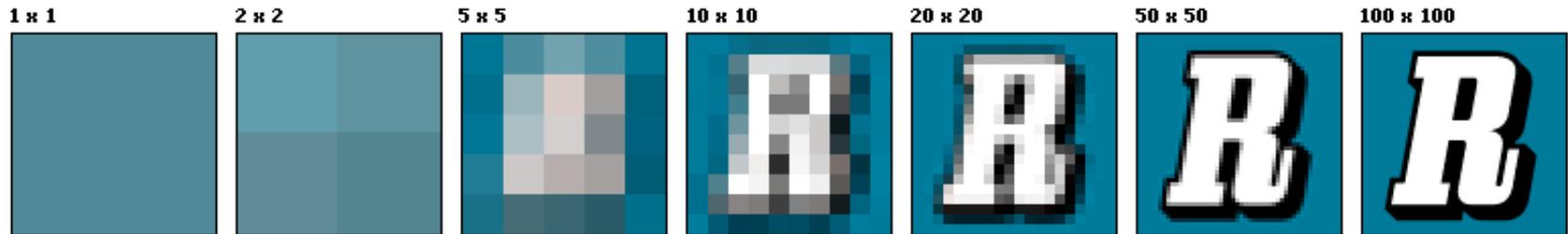
Pixel y resolución de imágenes



Pixel es la unidad mínima de una imagen digital y corresponde a un cuadrado de un solo color. El siguiente ejemplo corresponde a una fotografía de 1280 x 960 pix procesada con software ACDsee y mostrada con sucesivas ampliaciones hasta llegar a visualizar las unidades mínimas.



Resolución en impreso y pantalla



- Web/PC • 72 - 96 dpi
- Impreso • 150 dpi
- Profesional • 300 o más dpi

- SVGA • 800x600px (4:3)
- XGA • 1024x768px (4:3)
- HD 720 • 1280x720px (16:9)
- SXGA • 1280x1024px (5:4)
- HD 1080 • 1920x1080px (16:9)

Mapa de bits versus vectorial



Vectorial

- Descripción matemática y exacta de cada elemento de la imagen.
- Las imágenes son difíciles de construir y no logran un detalle realístico.
- No pierden resolución nunca.

Bitmap

- Descripción de la imagen pixel por pixel.
- Usado en fotografías, puede guardar detalles muy precisos vistos a distancia.
- La imagen pierde resolución a medida que se acerca.

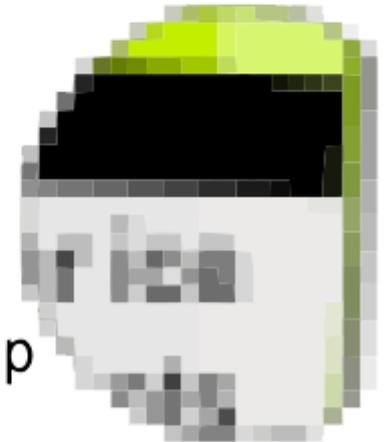


7x Magnification

Vector



Bitmap

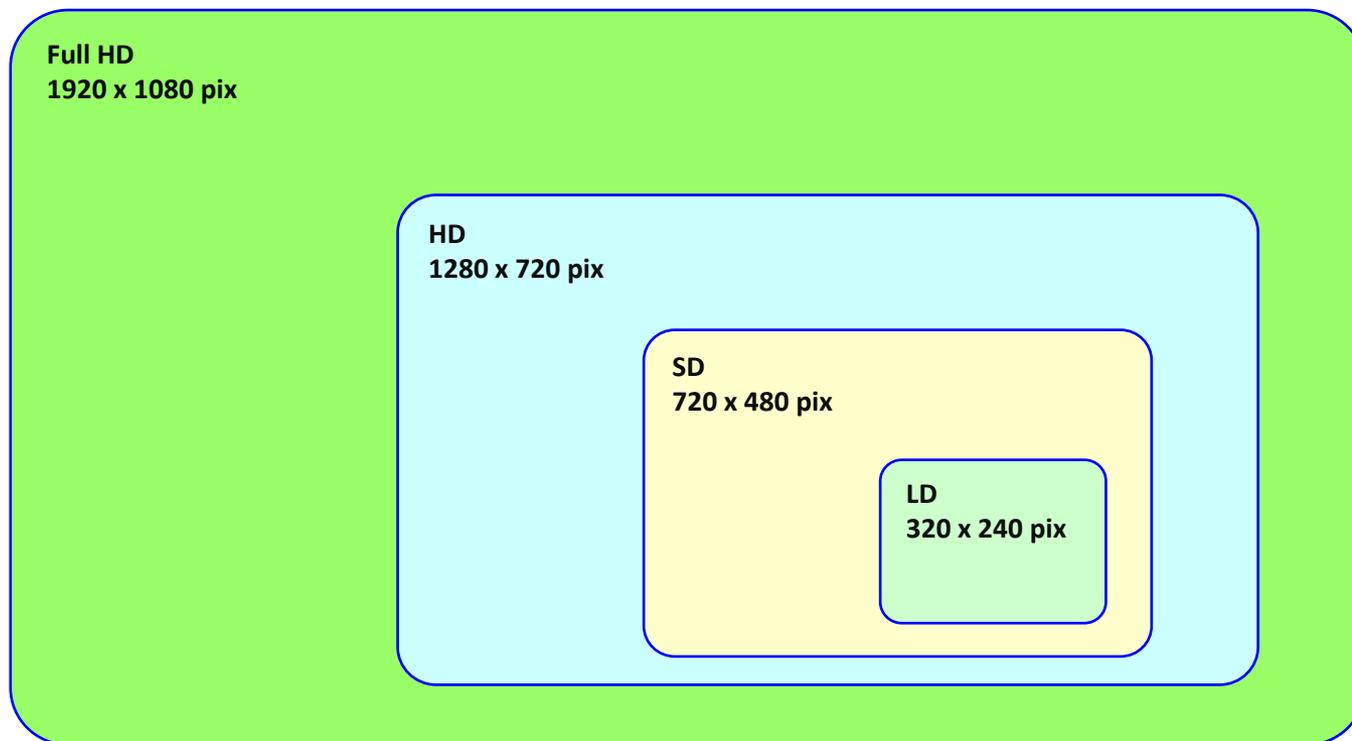


Resolución de pantalla



Dependiendo de la cantidad de píxeles horizontales y verticales se distinguen formatos de resolución independientes del tamaño de la pantalla.

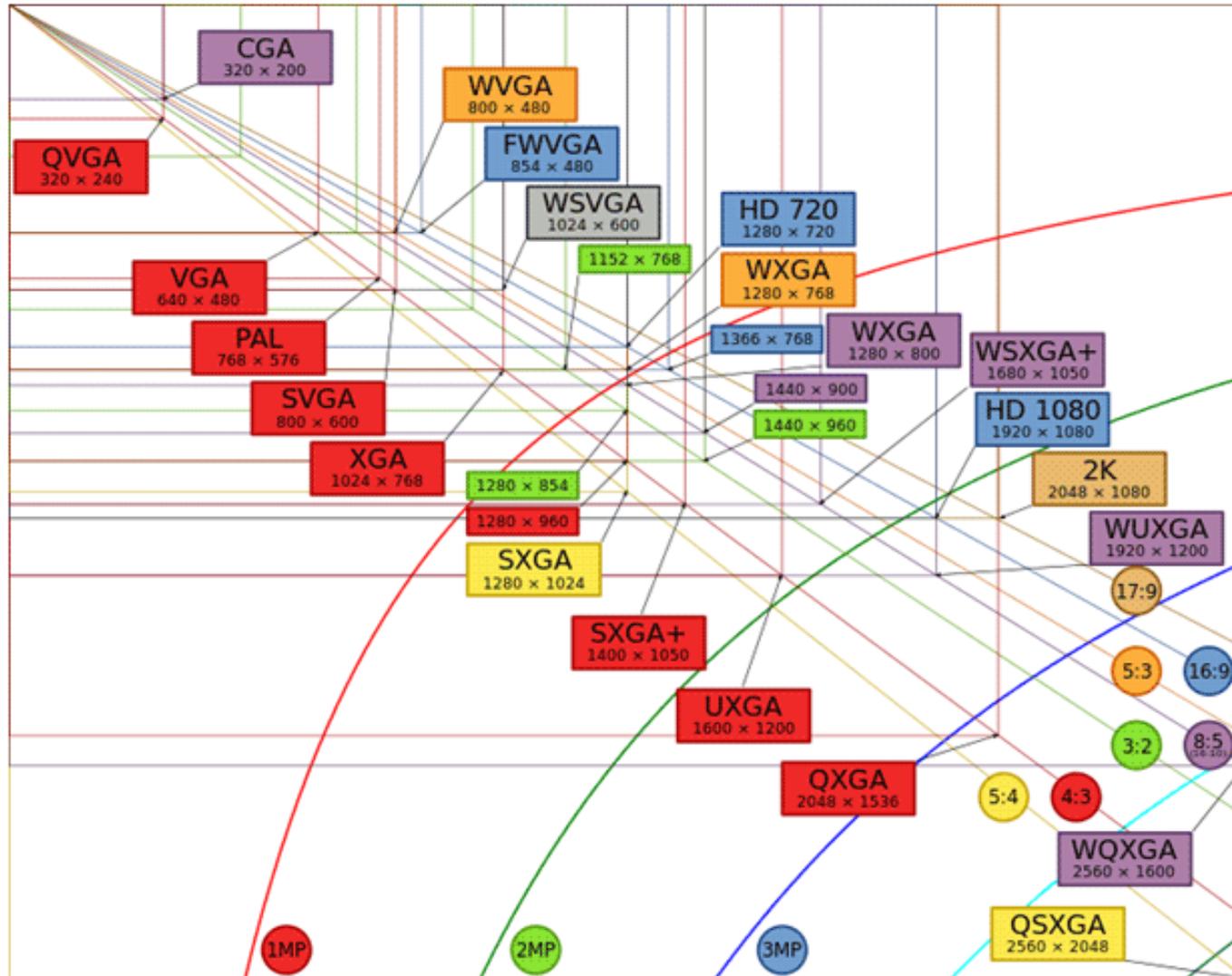
Designación	Descripción	Resolución (pix)	Calidad tipo
UHD	Ultra High Definition (Ultra Alta Definición)	3840x2160	UHDTV (4K)
Full HD	High Definition (Alta Definición)	1920x1080	HDTV (Blue Ray)
HD	High Definition (Alta Definición)	1280x720	HDTV
SD	Standard Definition (Definición Estándar)	720 x 480	DVD
LD	Low Definition (Baja Definición)	320 x 240	Celular



Resolución de pantalla



Cuadro general resolución de pantallas de televisión, computación, celulares, etc.



Fuente: http://commons.wikimedia.org/wiki/File:Vector_Video_Standards2.svg

Formatos de pantalla en televisión



4:3

16:9

21:9?



reducido



Recortado



Comprimido



Formato preferido por todos los estándares de televisión digital es el 16:9, no obstante, incluyen el 4:3.

Computación, videovigilancia pueden utilizar otros formatos de pantalla.

HDTV (Alta Definición)



Se refiere a mejorar la resolución del video mediante una pantalla que permita desplegar 1920x1080 pixeles (Full HD) y la fidelidad del audio a 5 canales (Dolby 5.1) mediante un Home Theater



HDMI : Conectores especiales para alta definición



Dolby 5.1 : Sistema de audio con sonido envolvente



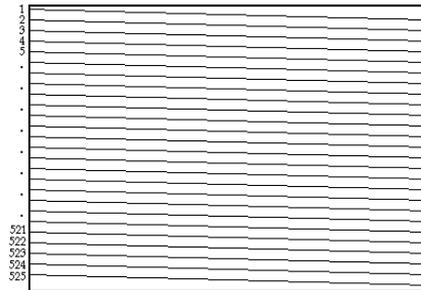
Pantalla de dimensiones panorámicas y sala con adecuación acústica brindan la mayor calidad de audio y video

La formación de la imagen en TV y computación

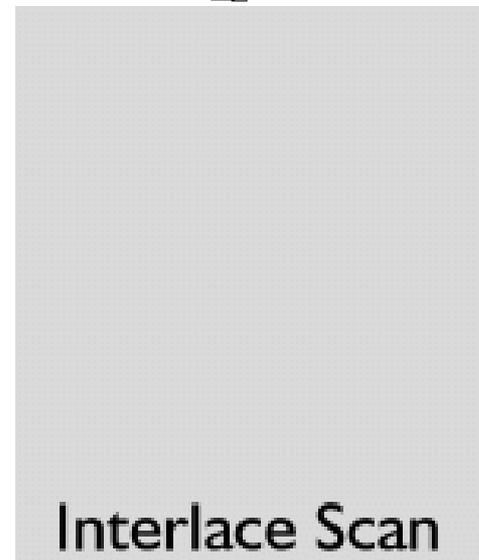
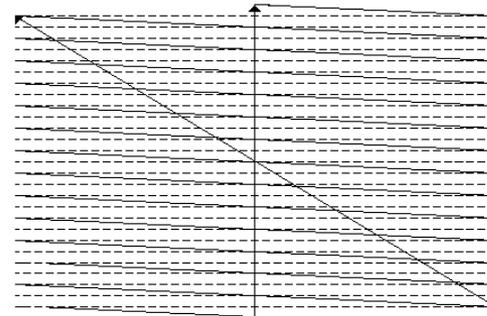


En un televisor convencional para formar un cuadro de imagen, un haz recorre la pantalla de arriba abajo y de izquierda a derecha. Este proceso se denomina barrido. Existen dos tipos de barrido:

barrido progresivo (p)



barrido entrelazado (i)



1: Principios físicos de imagen y sonido y fisiológicos de audición y visión humana y su conversión para el procesamiento y transmisión

2: Procesamiento y transmisión analógico y digital de señales de voz, audio, imagen y video

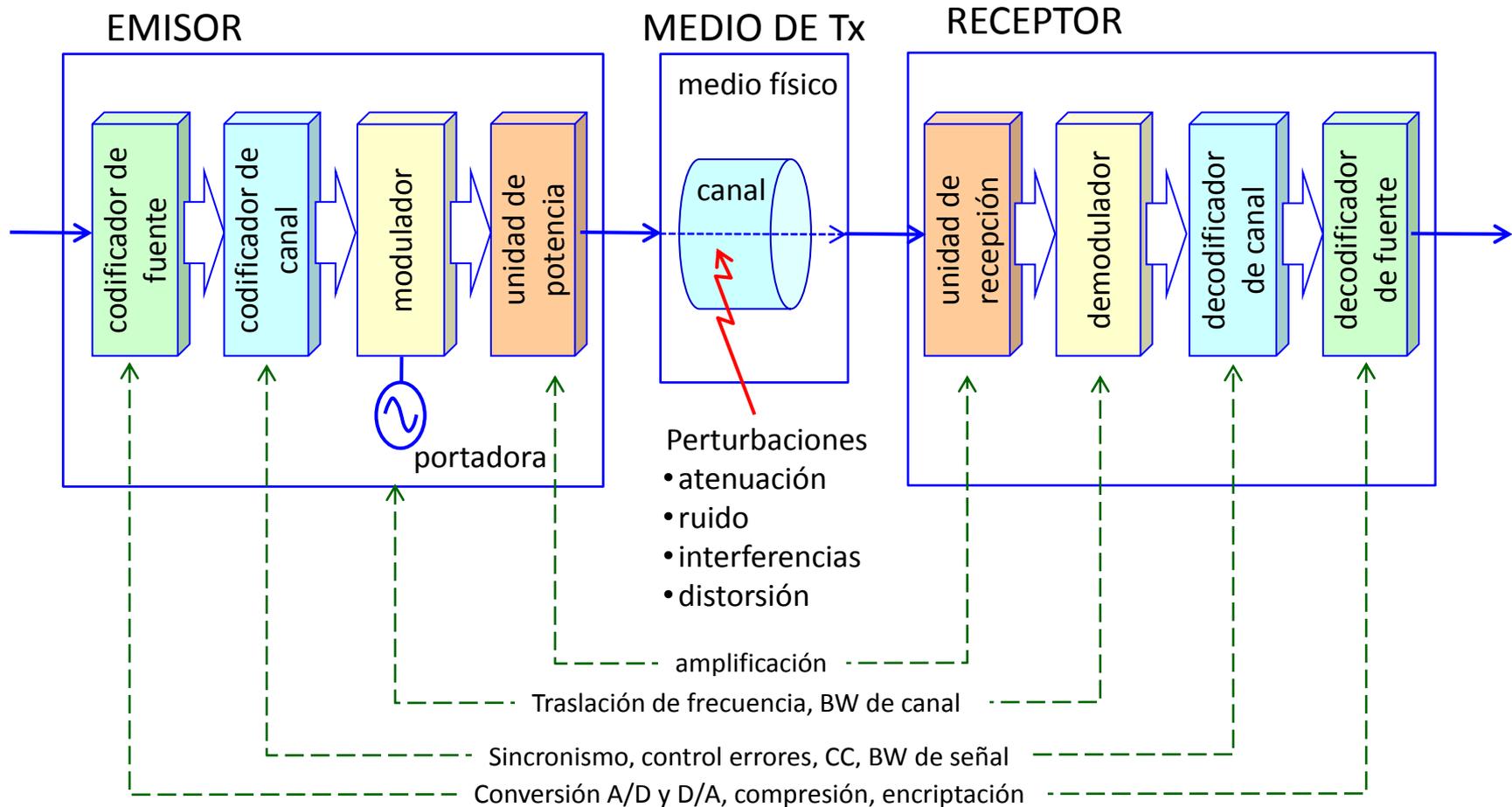
3: Características y requerimientos de una red IP para el transporte de señales de voz, audio y video

4: Estándares y normativa de sistemas de procesamiento y transmisión de voz y video IP

Procesamiento de la información en un sistema de telecomunicaciones



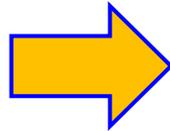
El transporte de señales de información por los medios de transmisión requiere codificarla para asegurar su integridad y recuperación en el receptor.



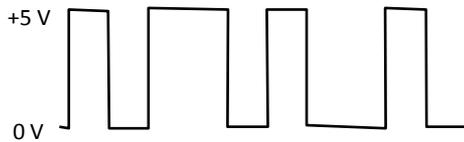
El problema fundamental de la transmisión de señales de información en la forma como se producen (banda base):



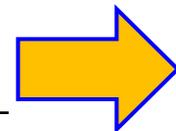
Analógica



- El ruido se incorpora a la señal de información
- Distorsión
- No admite procesamiento matemático

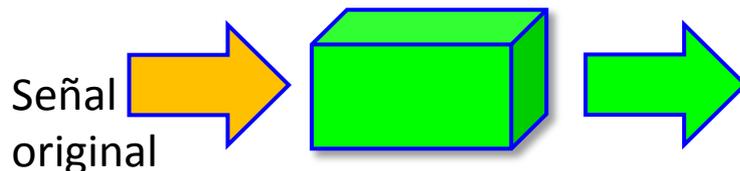


Digital (binario)



- Uso ineficiente del ancho de banda
- Distribución de energía no optimizada
- Contiene componente de CC
- Factible de pérdida de sincronismo

Por lo tanto, junto con digitalizar la señal de información es necesario efectuar la codificación de canal para adaptar la señal al medio de transmisión:

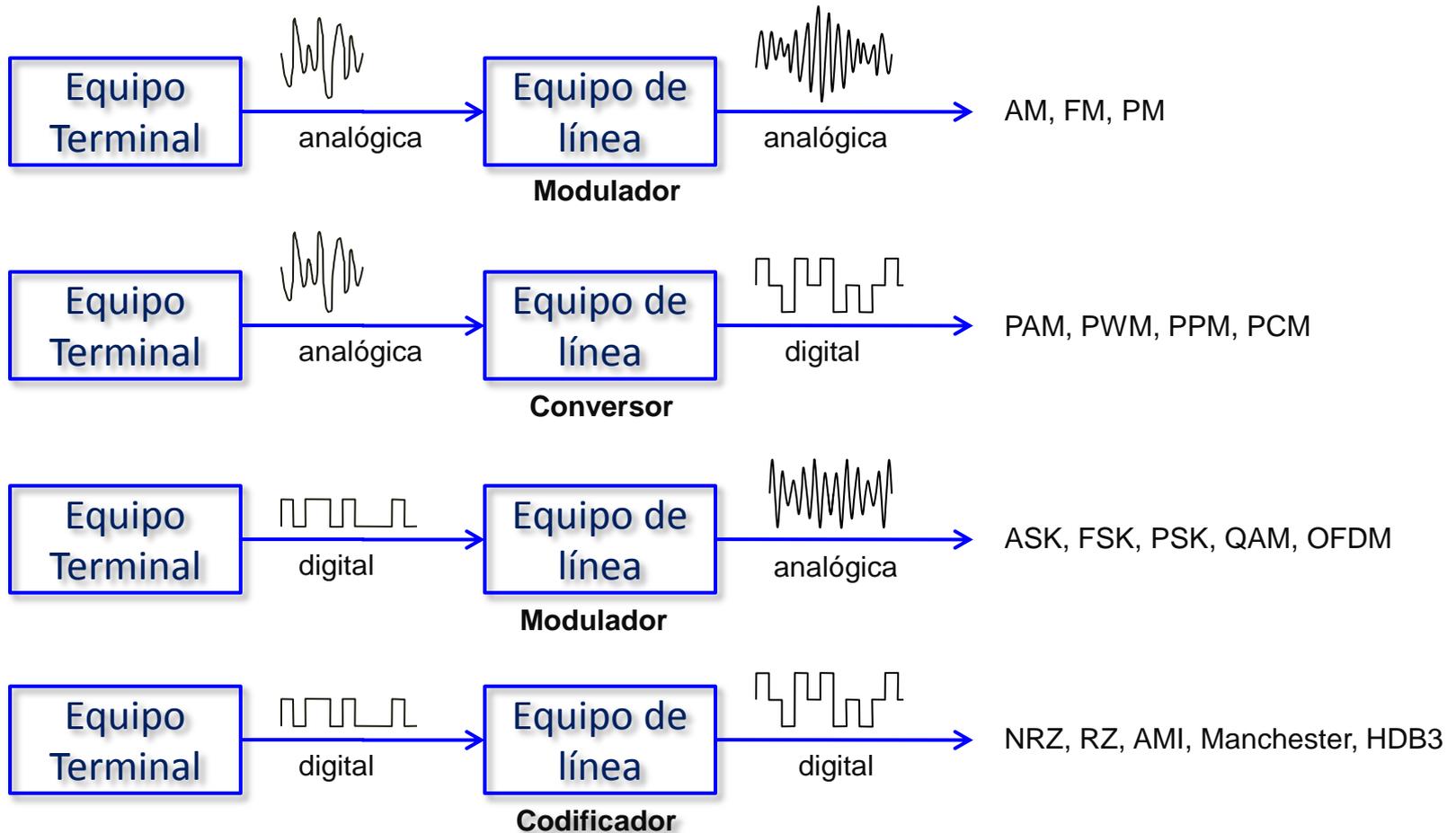


- Manejo del BW
- Distribución optimizada de la potencia en el BW
- Control del sincronismo
- Sin componente de CC

Procesamiento de la señal para la transmisión



Tanto la fuente de información como la señal de transmisión puede ser analógica o digital



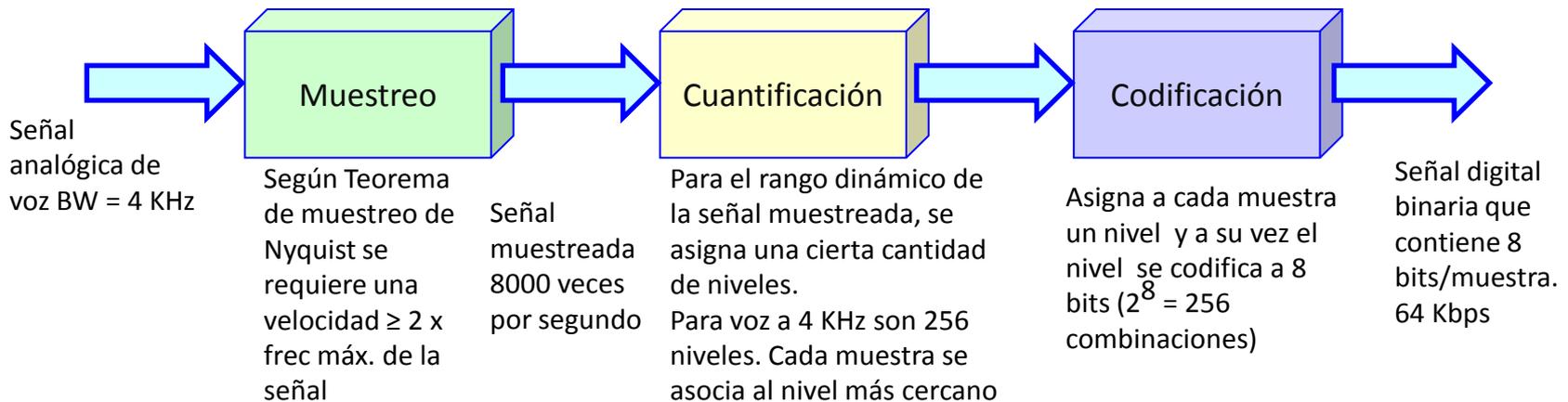
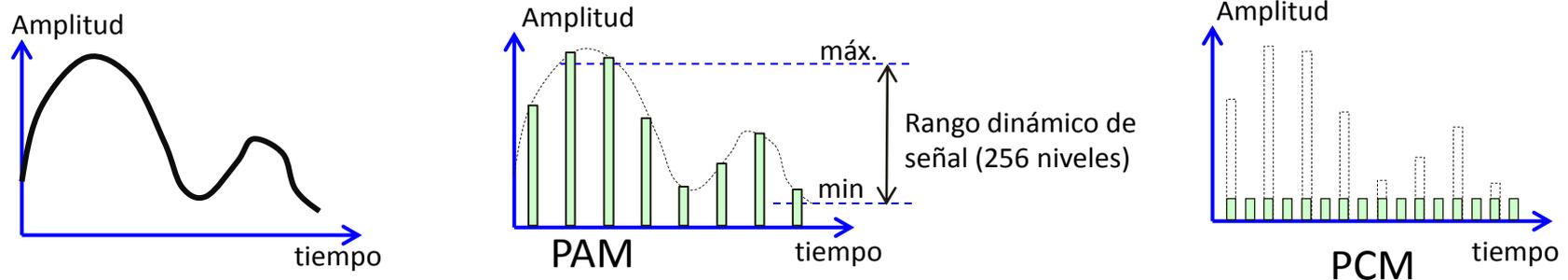


- El primer paso en VoIP (Voz sobre IP) es convertir la señal analógica en una señal digital.
- Desde hace muchos años se conoce el principio de PCM, el cual comienza hacia los años 40.
- Hacia los años 50 se comienza a trabajar sobre lo que se toma como el precursor del Vocoder, el cual progresa recién en la era digital, dado que la tecnología analógica no era suficiente.

Conversión análogo-digital de la voz



Estándar base [G.711 de ITU](#), define para la señal analógica de voz de un canal telefónico (BW= 4 KHz para una comunicación inteligible) un proceso de digitalización basada en 3 etapas: muestreo, cuantificación y codificación.



$$\text{Velocidad binaria} = 8000 \text{ muestras/seg} \times 8 \text{ bits/muestra} = 64 \text{ Kbps}$$

PAM: Pulse Amplitud Modulation

PCM: Pulse Coded Modulation

- La ITU normalizó las distintas técnicas de compresión, algunas de las cuales utilizaremos en VoIP.
- De lo visto anteriormente, la ITU, generó una serie de especificaciones para el servicio de “Multimedia” según el medio de transporte, especificando además que tipo de compresión utilizar en cada caso.

Año	Norma	Técnica	Velocidad (Kb/s)
1972	G.711	PCM	64
1984	G.721	ADPCM	32
1988	G.722	ADPCM	48, 56 Y 64
1992	G.728	LD-CELP	16
1995	G.729	CS-CELP	8
1995	G.723.1	MPC-MLQ	5,3 Y 6,4

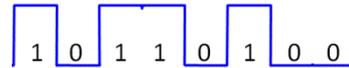
Generación de bits de imagen de video y compresión



¿Qué ocurriría si los bits de video se almacenaran tal como son producidos?



Cámara digital HD



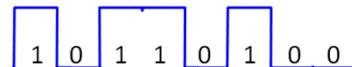
Disco Duro de 160 GB
se llenaría en
14 minutos y 17 segundos

$$(30 \text{ cuadros/seg}) \times (1920 \times 1080 \text{ pixel/cuadro}) \times (24 \text{ bits/pixel}) \\ = 1.493 \times 10^6 \text{ bps} + \text{el audio digital... (más de 1,49 Gbps)}$$

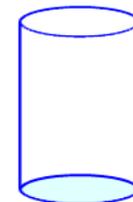
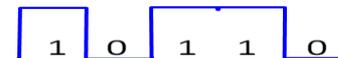
Comprimir significa reducir la cantidad de bits hasta un nivel razonable que no degrade la calidad de la imagen. Entre las técnicas más eficientes se encuentra MPEG en sus versiones MPEG-2, MPEG-4 y la H.264



Cámara digital HD



> 1,49 Gbps



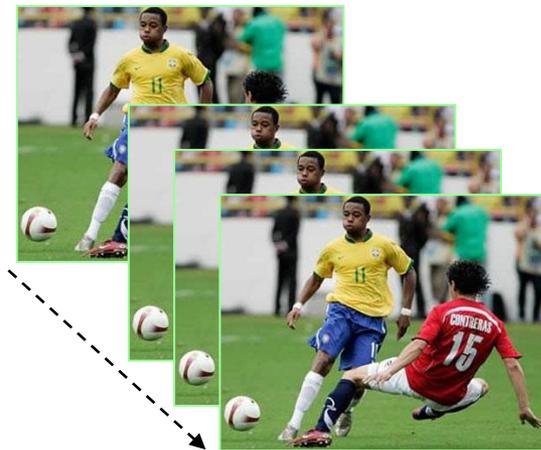
Disco Duro de 160 GB
se llenaría en
35 horas y 33 minutos

Dependiendo del tipo de imagen podría ser 10 Mbps para una señal HD

Compresión de video

Objetivos de la compresión: Reducción del tamaño de la información para facilitar su almacenamiento y reducción de costos de transmisión, mediante:

- ✓ Eliminación de información redundante (sin pérdidas). Esto se denomina compresión espacial y temporal.
- ✓ Eliminación de información no redundante que afecta marginalmente la información (con pérdidas). Esto se denomina submuestreo.



4: 4: 4

Ejemplo: Requerimiento de capacidad y velocidad para NTSC sin compresión:

Resolución de imagen

$720 \times 480 \text{ pix/fotograma} = 345.600 \text{ pixeles}$
 $< 0,5 \text{ Megapixeles}$

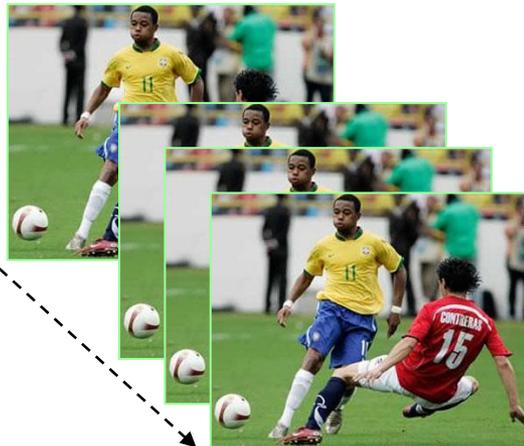
Exigencia al canal de transferencia

$720 \times 480 \text{ pix/fotograma} \times 8 \text{ bits/pixel} \times 3$
 $\text{colores/pixel} \times 30 \text{ fps} = 248,8 \text{ Mbps}$

Para la compresión de vídeo se aplican dos técnicas:

- ✓ Compresión espacial o intraframe: se aprovecha la redundancia de información que hay en la imagen de cada fotograma, como en la imágenes JPEG y JPEG Motion.
- ✓ Compresión temporal o interframe: se aprovecha la redundancia de información que hay entre fotogramas consecutivos.

La compresión interframe siempre lleva incluida la intraframe.

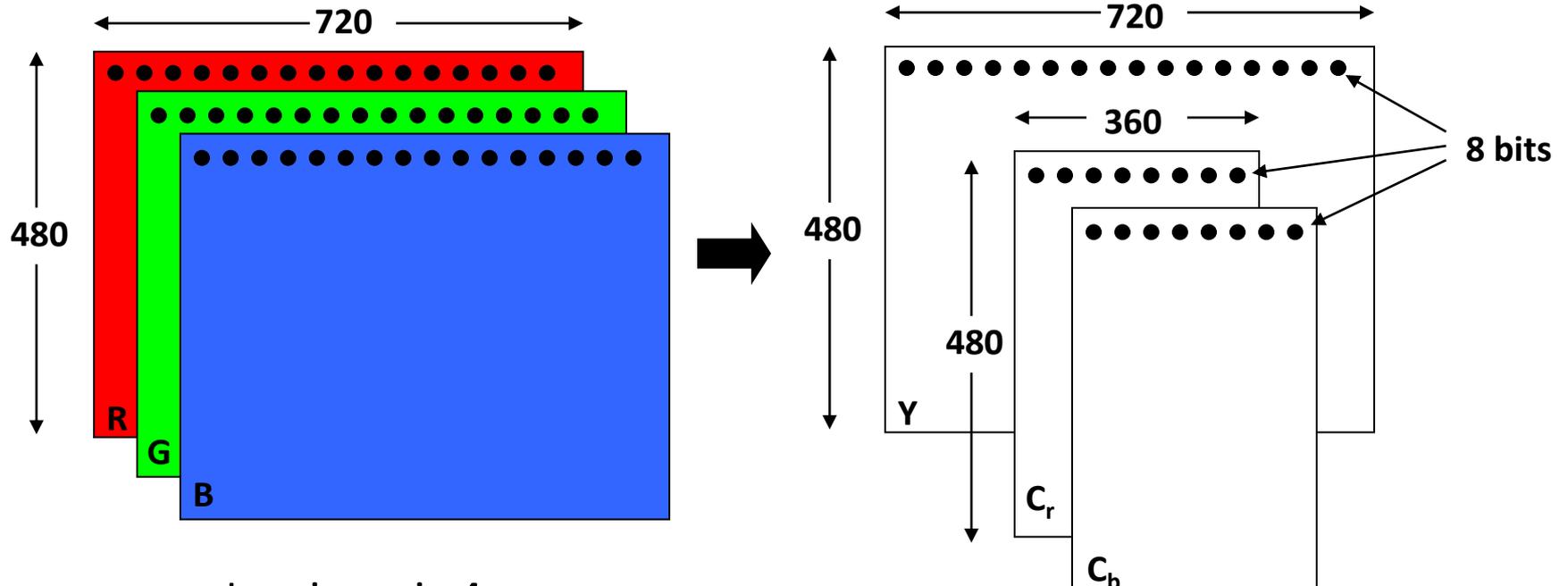


Submuestreo (compresión de vídeo con pérdidas)



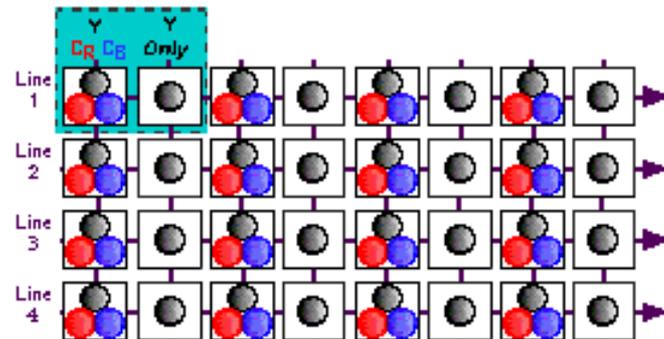
- ✓ La reducción de la resolución en las componentes de crominancia se denomina **submuestreo** (subsampling).
- ✓ El submuestreo se basa en la menor sensibilidad del ojo humano a la crominancia.
- ✓ El submuestreo 4:2:2 de CCIR-601 reduce la información de crominancia a la mitad (sin submuestreo el caudal total sería 248,832 Mbps).
- ✓ La información de crominancia puede reducirse aún más (a la cuarta parte) aplicando submuestreo 4:1:1 o 4:2:0.
- ✓ Este submuestreo degrada un poco la calidad de color (implica pérdida de información), pero la diferencia con 4:2:2 es pequeña y sólo suele ser percibida por profesionales o en situaciones extremas.

Submuestreo 4:2:2

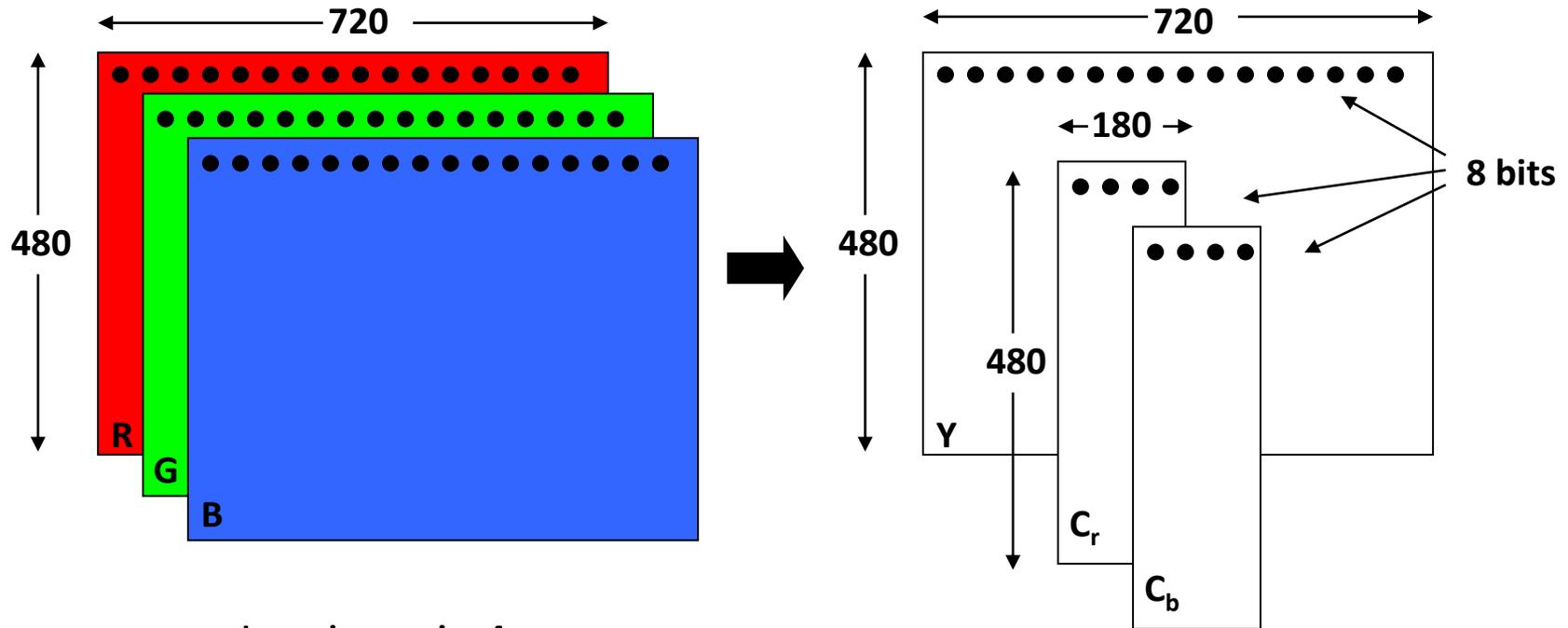


Luminancia 4
Crominancia 2+2

$$(720 \times 480 + 360 \times 480 + 360 \times 480) \times 8 \text{ bits} = 5.529.600 \text{ bits/cuadro}$$

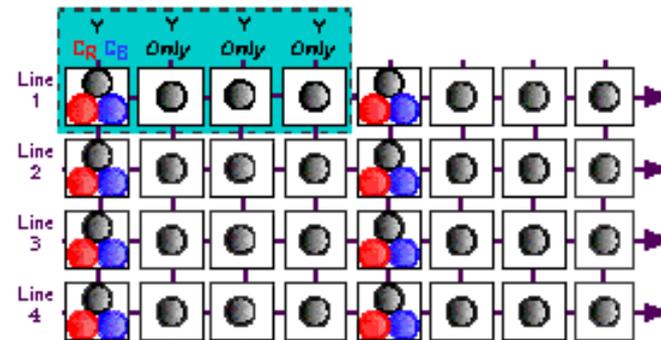


Submuestreo 4:1:1

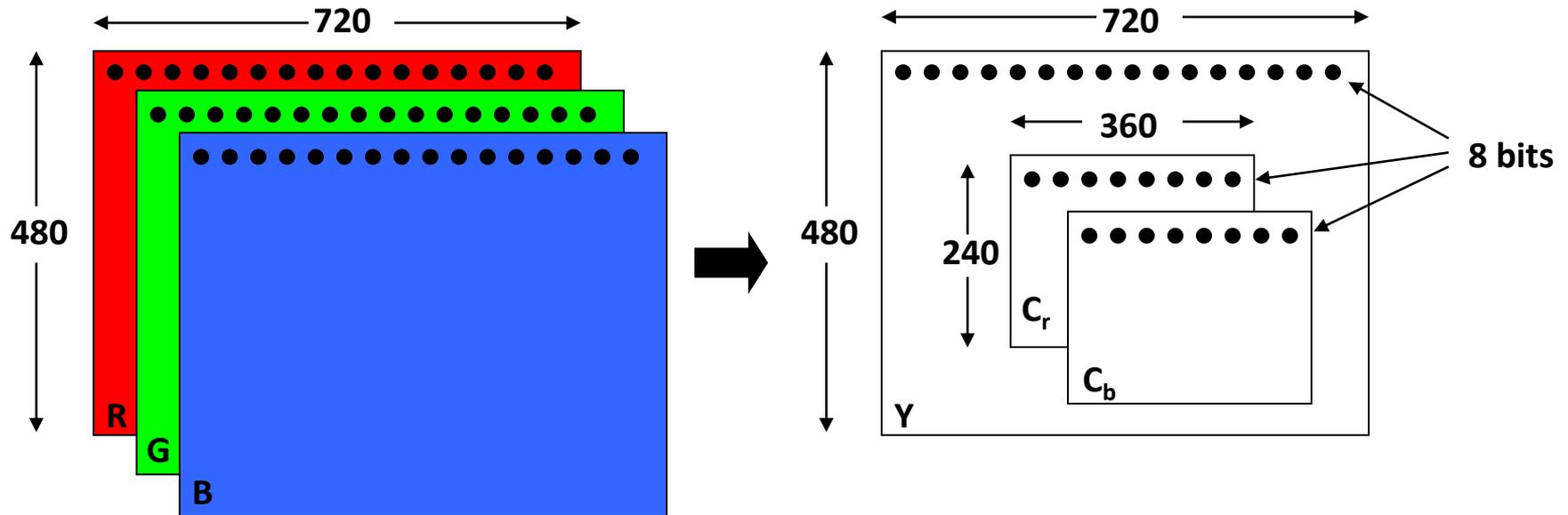


Luminancia 4
Crominancia 1+1

$$(720 \times 480 + 180 \times 480 + 180 \times 480) \times 8 \text{ bits} = 4.147.200 \text{ bits/cuadro}$$

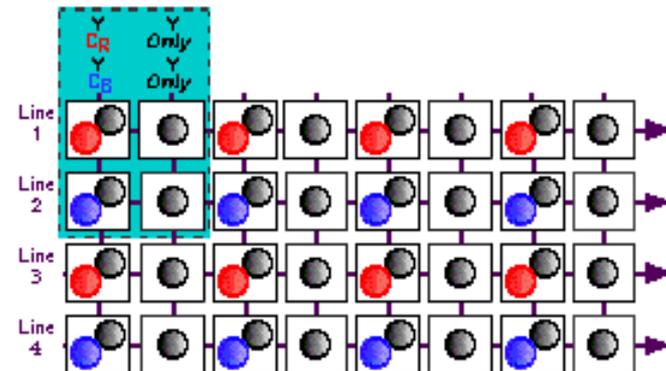


Submuestreo 4:2:0



Luminancia 4
Crominancia 2+0

$$(720 \times 480 + 360 \times 240 + 360 \times 240) \times 8 \text{ bits} = 4.147.200 \text{ bits/cuadro}$$



Formatos compresión de vídeo



Sistema	Compresión Espacial (DCT)	Compresión temporal	Complejidad compresión	Eficiencia	Retardo
M-JPEG	Sí	No	Media	Baja	Muy pequeño
H.261	Sí	Limitada (fotog. I y P)	Elevada	Media	Pequeño
MPEG-1/2	Sí	Extensa (fotog. I, P y B)	Muy elevada	Alta	Grande
H.264 MPEG-4	Sí	Extensa (fotog. I, P y B)	Enorme	Alta	Medio Grande

Caudal requerido por los sistemas de compresión de vídeo más comunes



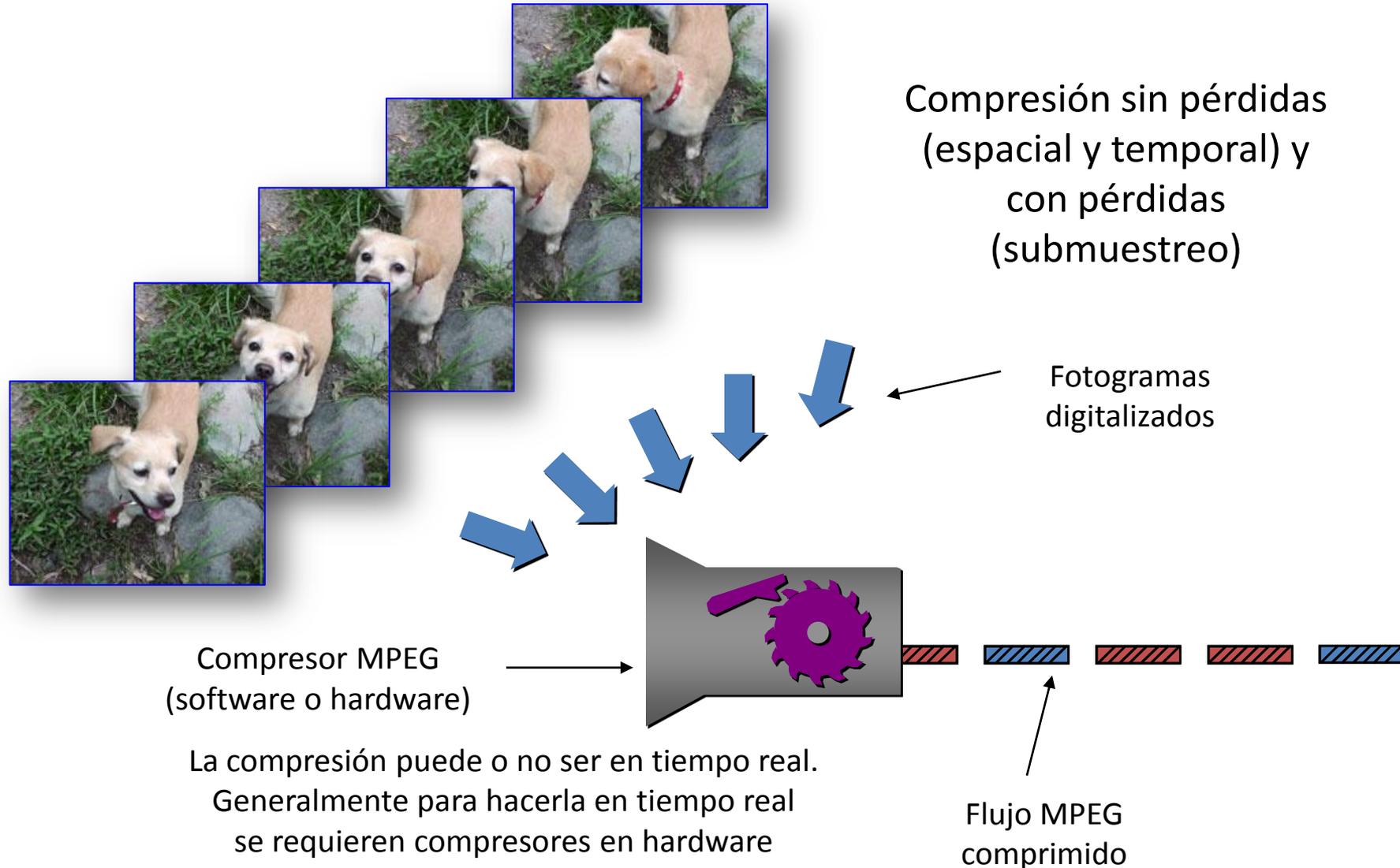
	Estándar/Formato	Ancho de banda típico	Ratio de compresión
Bajo retardo	CCIR 601	170 Mbps	1:1 (Referencia)
	M-JPEG	10-20 Mbps	7-27:1
	H.261	64 Kbps – 2000 Kbps	24:1
	H.264	28,8 - 768 Kbps	50:1
Elevado retardo	MPEG-1	0,4 - 2,0 Mbps	100:1
	MPEG-2	1,5 - 60 Mbps	30-100:1
	MPEG-4	28,8 - 500 Kbps	100-200:1

Vídeo M-JPEG (Motion JPEG)



- ✓ Es el más sencillo. Trata el vídeo como una secuencia de fotografías JPEG, sin aprovechar la redundancia entre fotogramas.
- ✓ Algoritmos DCT (Discrete Cosine Transform)
- ✓ Poco eficiente, pero bajo retardo.
- ✓ Usado en:
 - ✓ Algunos sistemas de grabación digital y de edición no lineal (edición independiente de cada fotograma)
 - ✓ Algunos sistemas de videoconferencia (bajo retardo).
- ✓ No incluye soporte estándar de audio. El audio ha de codificarse por algún otro sistema (p. Ej. CD-DA) y sincronizarse por mecanismos no estándar.

Funcionamiento de MPEG



Vídeo MPEG (MPEG-1)



- ✓ Submuestreo 4:2:0 (25% ahorro respecto 4:2:2)
- ✓ Dos formatos posibles:
 - SIF (Standard Interchange Format): en PAL Y: 352 x 288 pixels, C_r y C_b : 176 x 144 pixels
 - QSIF (Quarter SIF): Y: 176 x 144; C_r y C_b : 88 x 72
- ✓ Dos tipos de compresión (simultáneamente):
 - Espacial: como en JPEG
 - Temporal: se aprovecha la semejanza que cada fotograma tiene con los que le rodean.
- ✓ El primer fotograma se digitaliza como una imagen JPEG
- ✓ De los siguientes fotogramas sólo se digitalizan los cambios respecto al anterior.
Para localizar los cambios:
 - Se 'cuadrícula' la imagen en macrobloques, cada uno formado por 16x16 pixels de Y (8x8 de C_r y 8x8 de C_b)
 - Si se detecta que un macrobloque ha cambiado de posición esto se indica mediante un vector de movimiento.
- ✓ Una imagen SIF (352x288) está formada por: $352/16 \times 288/16 = 22 \times 18 = 396$ macrobloques

Vídeo MPEG (tipos de fotogramas)



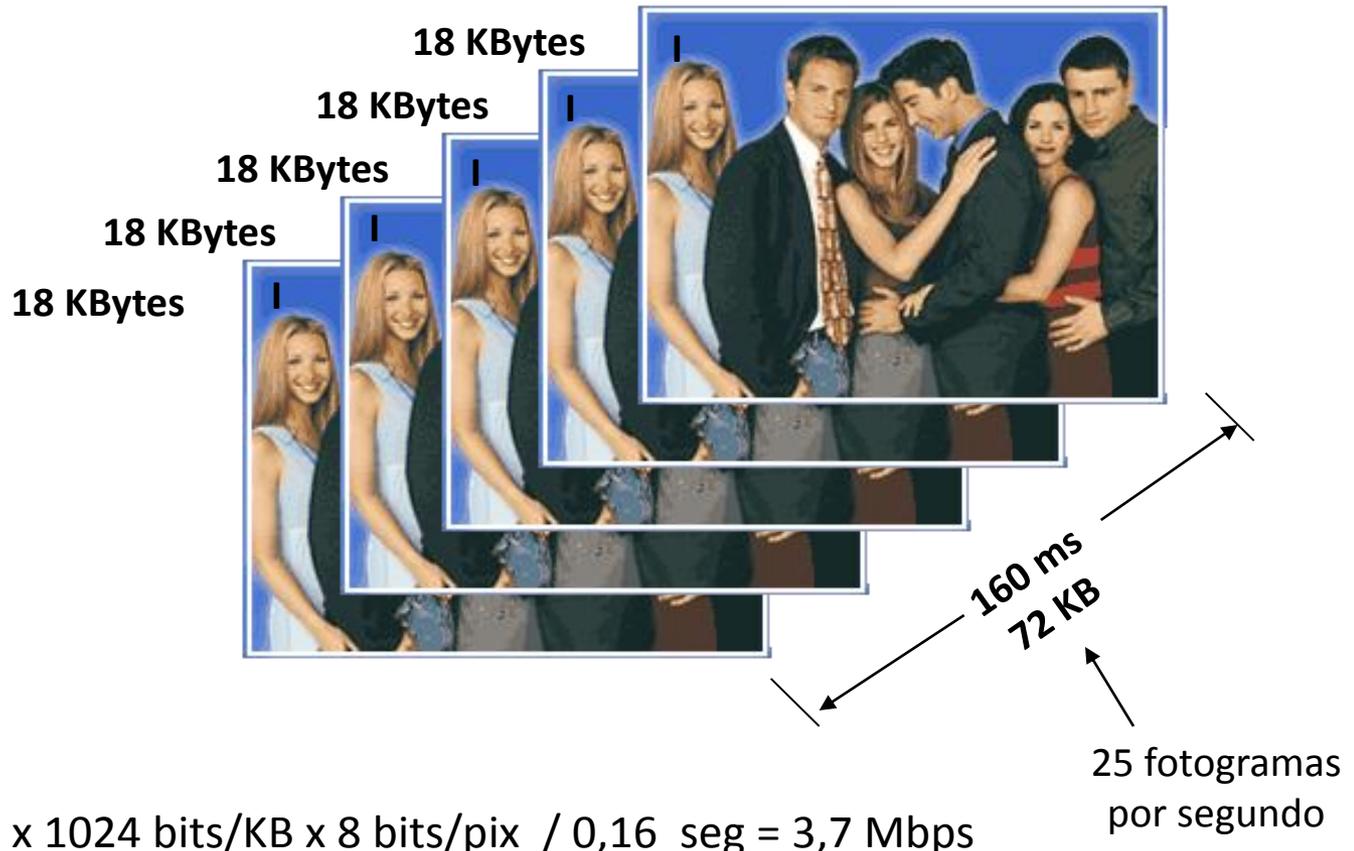
En MPEG se distinguen los siguientes tipos de fotogramas:

- ✓ **I (Intra)**: autocontenidos, solo compresión espacial (como JPEG)
- ✓ **P (Predictive)**: referido al P/I anterior. Compresión temporal por extrapolación mediante macrobloques. Un macrobloque pueden ser:
 - Inalterado: no modificado respecto al fotograma de referencia
 - Desplazado: (p. ej. un balón en movimiento) se describe por un vector de movimiento y eventualmente una corrección (diferencia respecto al original)
 - Nuevo: (p. ej. Lo que aparece detrás de una puerta que se abre) se describe por compresión espacial (como un fotograma I)
- ✓ **B (Bidireccional)**: compresión temporal con interpolación; referido al P/I anterior y al P/I posterior. Máxima compresión, máxima complejidad de cálculo. Suaviza la imagen, reduce el ruido.

Fotogramas I (Intra)



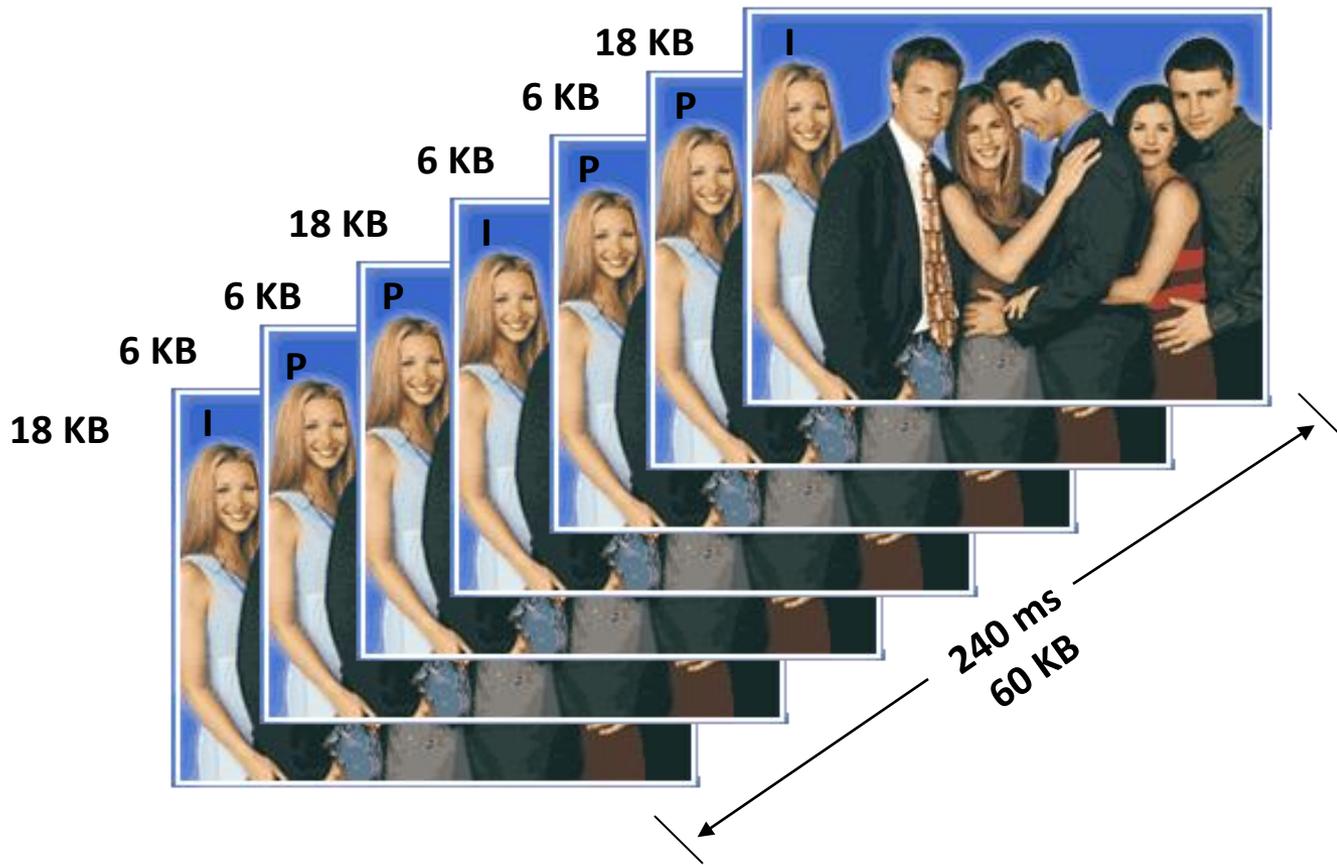
Los fotogramas Intra se codifican de forma autocontenida, sin referirse a otros fotogramas



Fotogramas P (Predictivos)



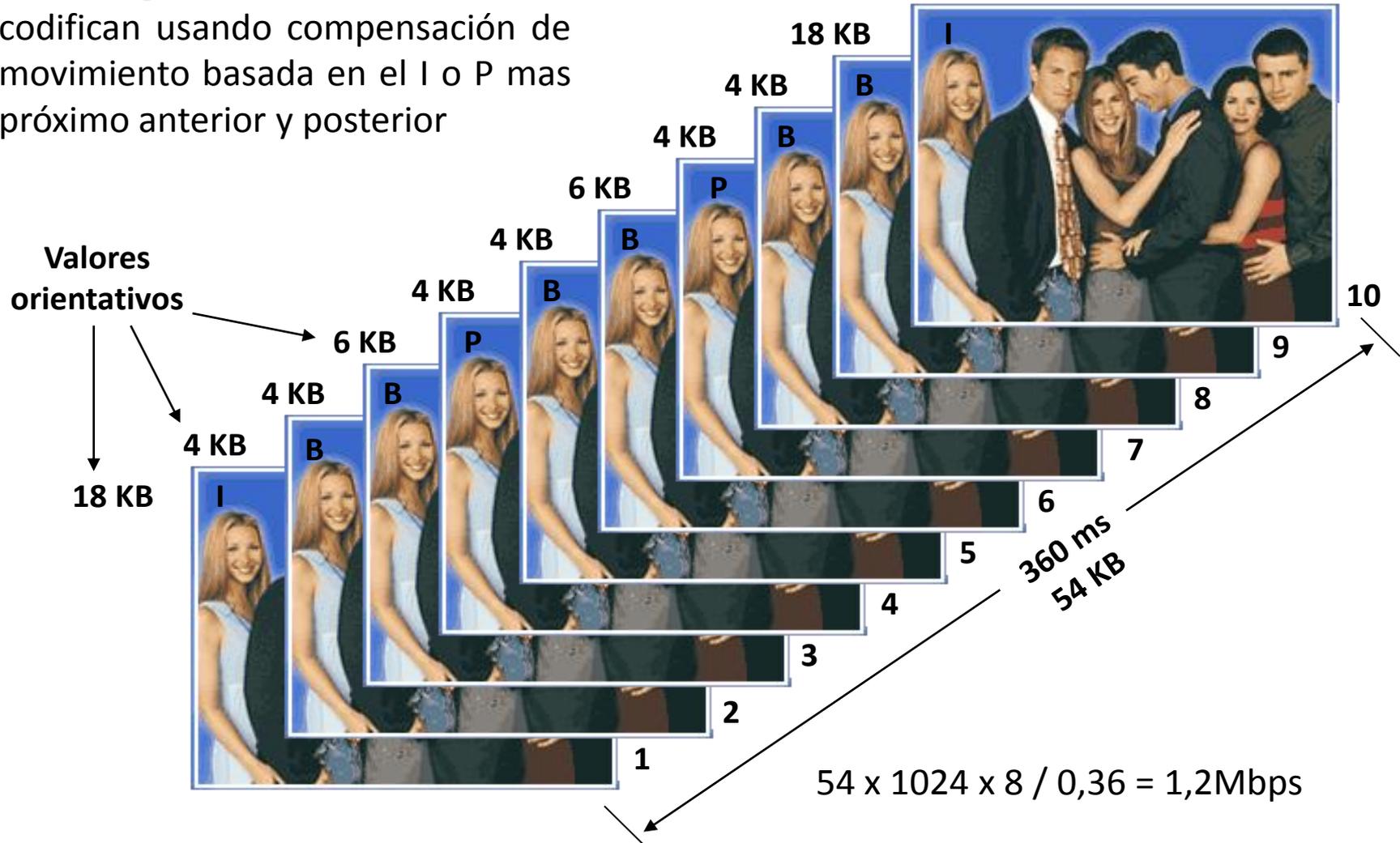
Los fotogramas Predictivos se codifican usando compensación de movimiento basada en el fotograma I o P anterior



$$60 \times 1024 \times 8 / 0,24 = 2,0 \text{ Mbps}$$

Fotogramas B (Bidireccionales)

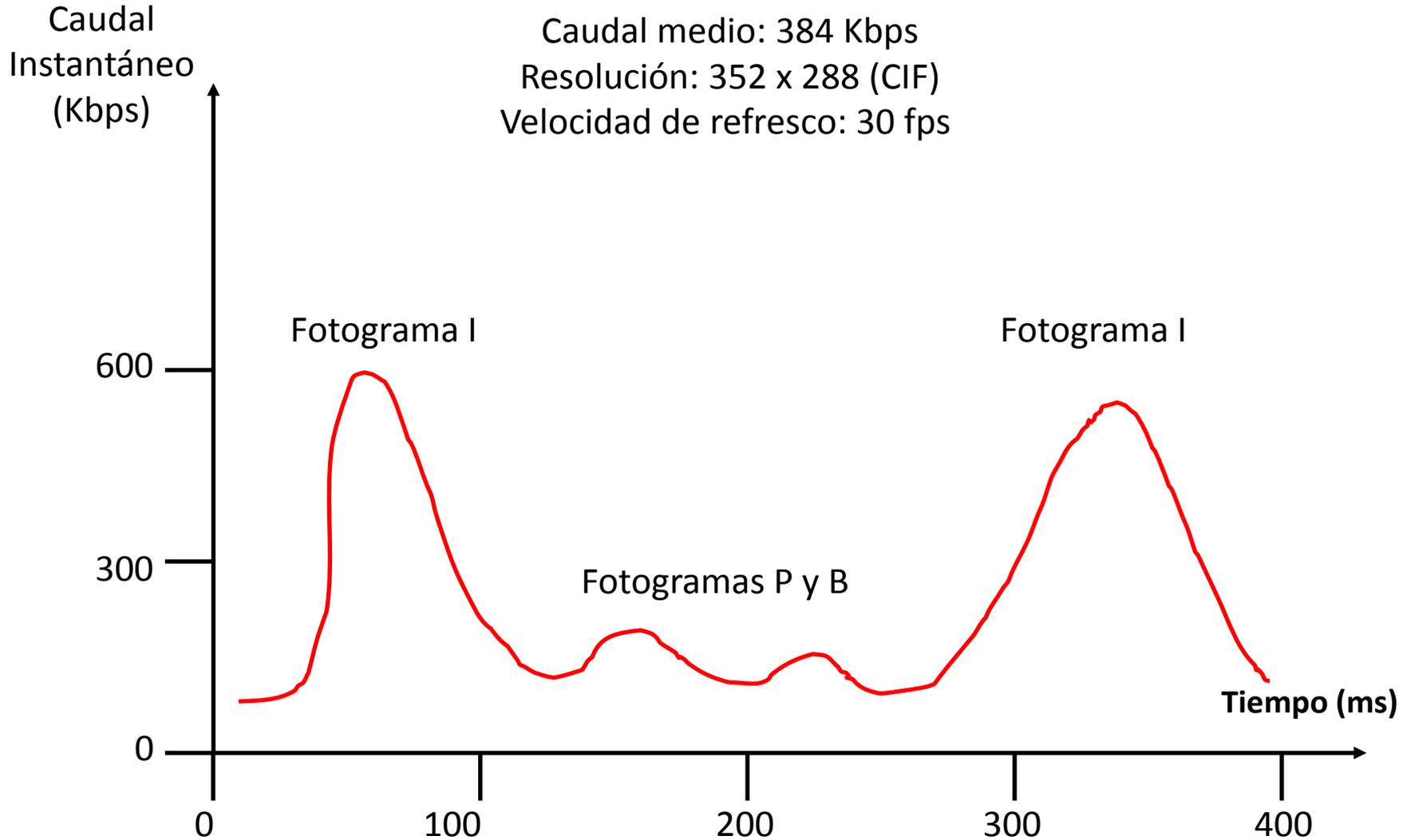
Los fotogramas Bidireccionales se codifican usando compensación de movimiento basada en el I o P mas próximo anterior y posterior



$$54 \times 1024 \times 8 / 0,36 = 1,2\text{Mbps}$$

Orden de transmisión: 1,4,2,3,7,5,6,10,8,9,...

Caudal de una videoconferencia



Audio MPEG-1

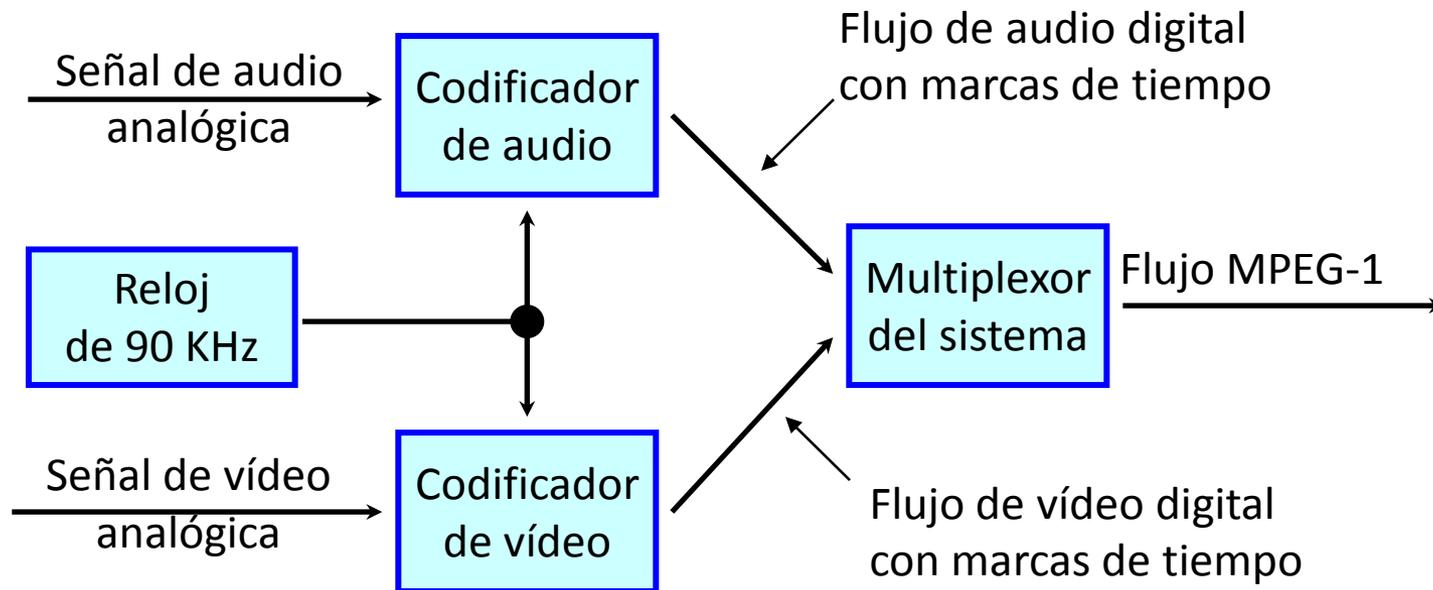
- ✓ Muestreo mono o estéreo a 32 KHz, 44.1 KHz (CD) o 48 KHz(DAT).
- ✓ Compresión psicoacústica (con pérdidas) asimétrica.
- ✓ De 32 a 448 Kbps por canal de audio
- ✓ Tres capas en orden ascendente de complejidad/calidad:
 - Capa I: buena calidad con 192-256 Kbps por canal; no se utiliza
 - Capa II: calidad CD con 96-128 Kbps por canal
 - Capa III: calidad CD con 64 Kbps por canal
- ✓ Cada capa incorpora nuevos algoritmos, y engloba los de las anteriores.
- ✓ **Capa III** usada en DAB (Digital Audio Broadcast) y en MP3

MPEG-1, capa 3 corresponde a mp3, el compresor de audio más difundido en el mundo.

Sincronización de audio y vídeo MPEG



Teniendo en cuenta que el flujo de video es generalmente de mayor volumen que el de audio, se hace necesario sincronizar los flujos de información visual con auditiva.



Durante la decodificación se realiza el proceso inverso

Vídeo MPEG-2 (1)



- ✓ Extensión compatible de MPEG-1
- ✓ Diseñado para televisión digital:
 - Optimizado para transmisión, no almacenamiento
 - Prevé vídeo entrelazado (TV) además de progresivo (MPEG-1 era sólo progresivo)
- ✓ Según los valores de los parámetros de muestreo utilizados se definen en MPEG-2 cuatro **niveles**:
 - **Bajo:** 352 x 288 (compatible MPEG-1)
 - **Principal:** 720 x 480 (equivalente CCIR 601 para NTSC)
 - **Alto-1440:** 1440 x 1152 (HDTV 4:3)
 - **Alto:** 1920 x 1080 (HDTV 16:9)



- ✓ Además de los niveles se definen seis **perfiles** según el submuestreo y algoritmo de compresión utilizado. Los perfiles posibles son:
 - **Simple:** para codecs de bajo costo
 - **Principal:** el más utilizado
 - **SNR**
 - **Espacial**
 - **Alto**
 - **4:2:2**
- } Para alta calidad
- ✓ No todas las combinaciones nivel-perfil están permitidas
 - ✓ Cada combinación tiene un caudal máximo previsto
 - ✓ TV digital y DVD utilizan nivel y perfil principal ML@MP (Main Level @ Main Profile)

Caudales de Niveles y Perfiles MPEG-2



Perfiles		Simple	Principal	SNR Escal.	Espacial Escal.	Alto	4:2:2 (Studio)
Submuestreo		4:2:0	4:2:0	4:2:0	4:2:0	4:2:0/2	4:2:0/2
Niveles	Alto 1920 x 1080 (HDTV 16:9)		80 Mbps			100 Mbps	
	Alto-1440 1440 x 1152 (HDTV 4:3)		60 Mbps		60 Mbps	80 Mbps	
	Principal 720 x 576 (CCIR 601)	15 Mbps	15 Mbps	15 Mbps		20 Mbps	50 Mbps
	Bajo 352 x 288 (MPEG1)		4 Mbps	4 Mbps			

Los mostrados son los caudales máximos previstos en el estándar para cada combinación de perfil y nivel.

Formatos de imagen



Maneras estandarizadas de guardar, comprimir y presentar información.

Formato	Descripción	Extensión
Tagged Image File Format (TIFF)	Formato no estándar y simple. Guarda los datos de manera muy lineal.	.tif, .tiff
Joint Pictures Expert Group (JPEG/JFIF)	Compresión basada en la percepción del ojo a los colores. Hay pérdida de resolución.	.jpg, .jpeg
Windows Bitmap (BMP)	Formato básico, pesado, para imágenes de alta resolución. Muy aceptado y usado.	.bmp
Graphics Interchange Format (GIF)	Imágenes con número de colores limitado a 256. Puede ser animado. Bueno para web.	.gif
Portable Networks Graphics (PNG)	Concebido para web. Imágenes de buena resolución y que aceptan transparencias.	.png
Standardized Vectors Graphics (SVG)	Formato vectorial. Estándar de la W3C para describir gráficos vectoriales. Basado en XML.	.svg

Formatos de video digital

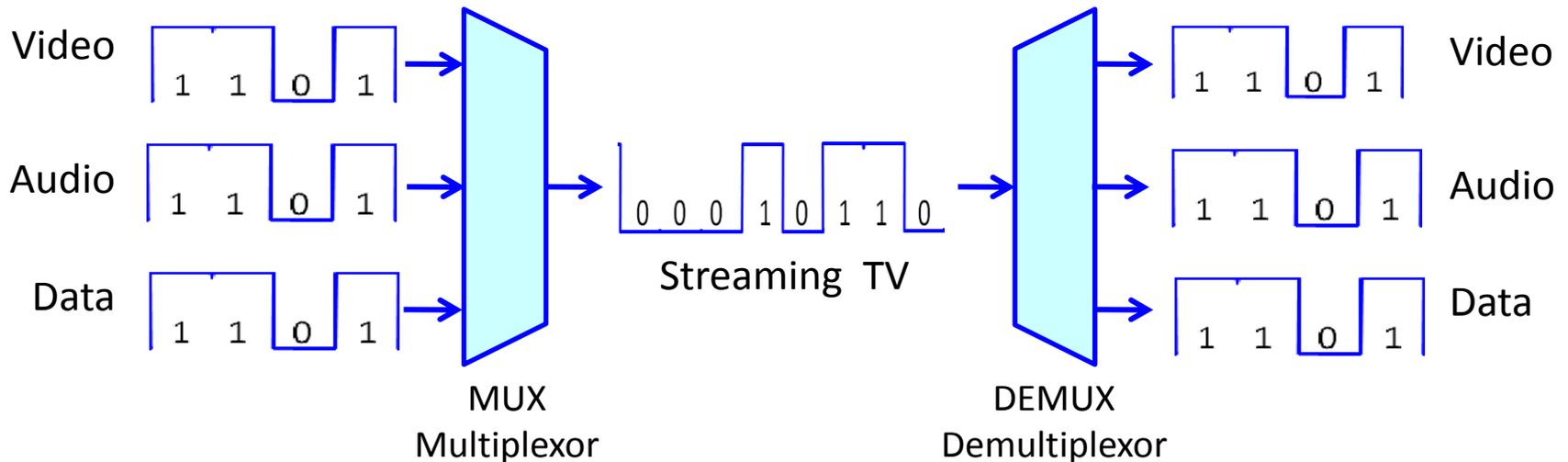


Formato	Descripción	Extensión
Audio/Video Interleave (AVI)	Desarrollado por Microsoft. Viejo. Video sin pérdida de calidad. Archivos pesados.	.avi
Motion Pictures Experts Group (MPEG)	Formato de compresión de audio y video desarrollado por el grupo con el mismo nombre. Ha tenido varias iteraciones. El famoso MP3 es la capa de audio de MPEG-1, formato que ya no es muy usado para video.	.mpg, .mpeg, .mp1, .mp2, .mp3, .m1v, .m1a, .m2a, .mpa, .mpv
Quicktime Movie (MOV)	Propiedad de Apple, junto con el códec para MPEG-4 usado en video HD actualmente.	.mov, .qt, .mp4
Windows Media Video (WMV)	Propiedad de Microsoft, contraparte del formato de audio similar, WMA.	.wmv, .asf
Flash Video (FLV)	Propiedad de Adobe, formato para web.	.flv
WebM	Propiedad de Google. Competidor de HTML5.	.webm
3G Partnership Project (3GPP)	Formato de video para celulares con tecnología de tercera generación (3G).	.3gp, .3g2

Multiplexación



Distintas señales de información se alinean en un solo flujo de alta velocidad y que pueden ser recuperadas en el otro extremo. Procesamiento MPEG considera esta función.



En televisión digital y videoconferencia se utiliza transmisión de múltiples señales (video, audio, datos) por cada canal y en forma simultánea también se pueden incorporar múltiples canales a un solo streaming mediante multiplexación.



El estándar MPEG establece dos posibilidades:

PS - Program Stream (Secuencia de Programa)

Válido para el almacenamiento o la transmisión de datos de un único programa sobre soportes o canales “seguros”:

- ✓ DVD
- ✓ CD
- ✓ HD

...

TS - Transport Stream (Secuencia de Transporte)

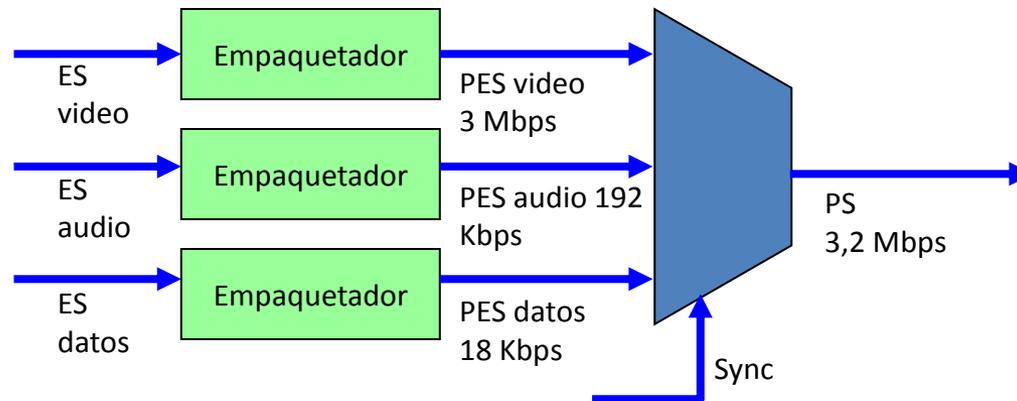
Diseñado para la transmisión o almacenamiento de los datos del programa en canales poco “seguros”, así como para el multiplexado de varios PS (Ejem. Satelital).

- ✓ Multiplexado
- ✓ Difusión
- ✓ Almacenamiento en cinta

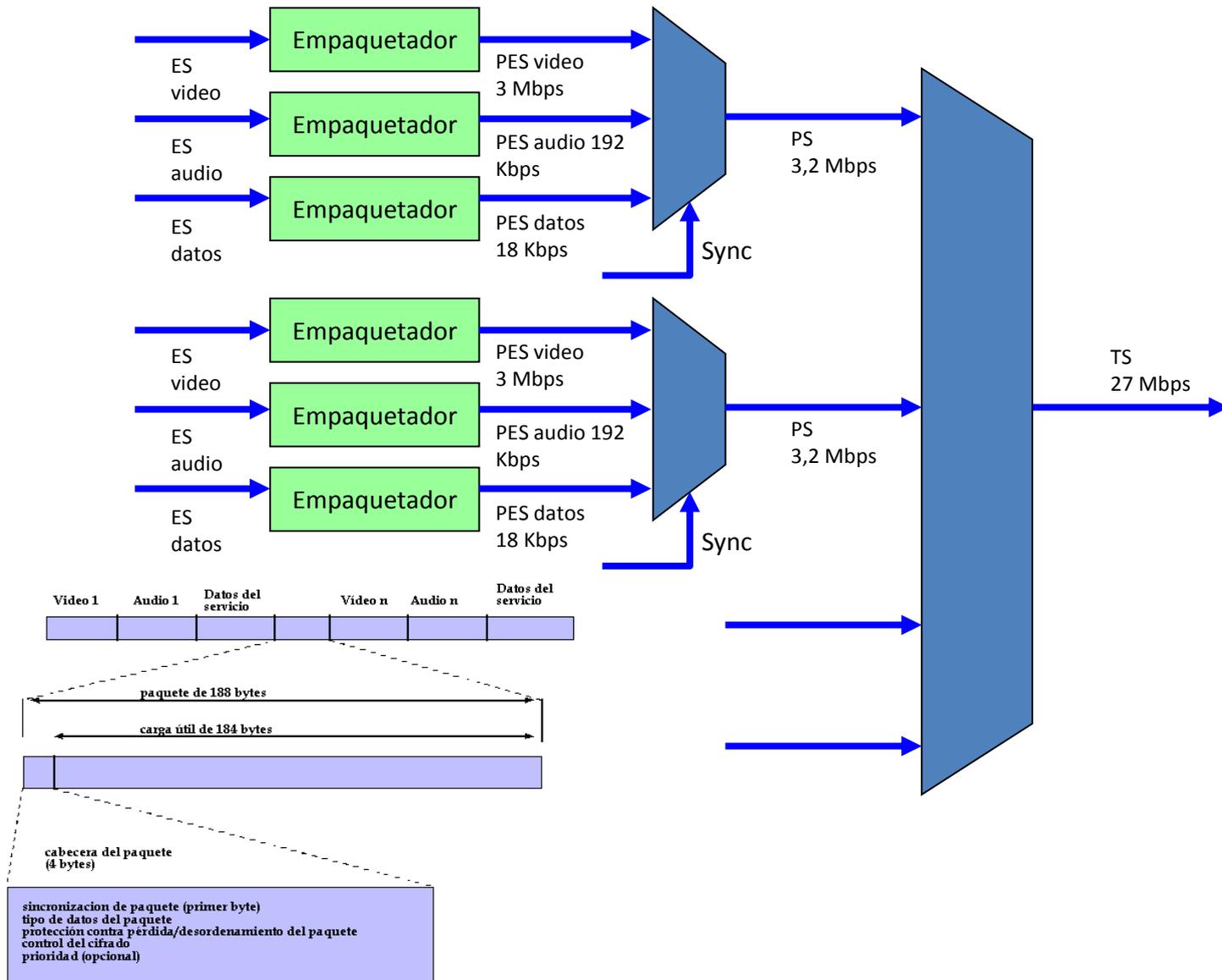
Formación del paquete MPEG (PS)

Cada elemento de los que van a constituir el PS (audio, video, datos...) se organiza en ES (Elementary Stream), secuencias elementales.

- ✓ Los datos de cada uno de dichos ES se empaquetan formando los PES (Packetized Elementary Stream)
- ✓ PS = PES Video + PES Audio + PES datos + Sincronismo



Formación de la trama MPEG-TS



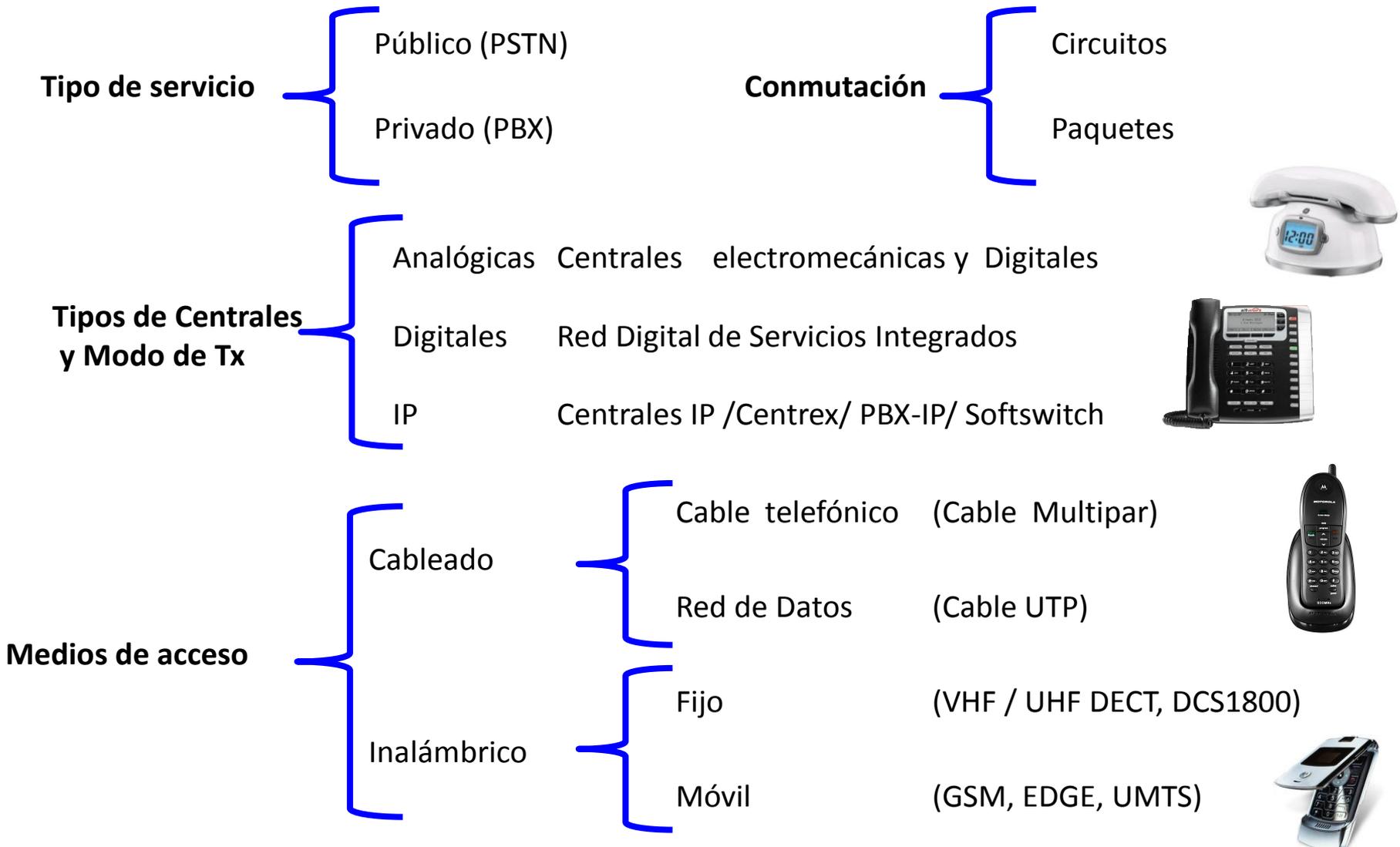
1: Principios físicos de imagen y sonido y fisiológicos de audición y visión humana y su conversión para el procesamiento y transmisión

2: Procesamiento y transmisión analógico y digital de señales de voz, audio, imagen y video

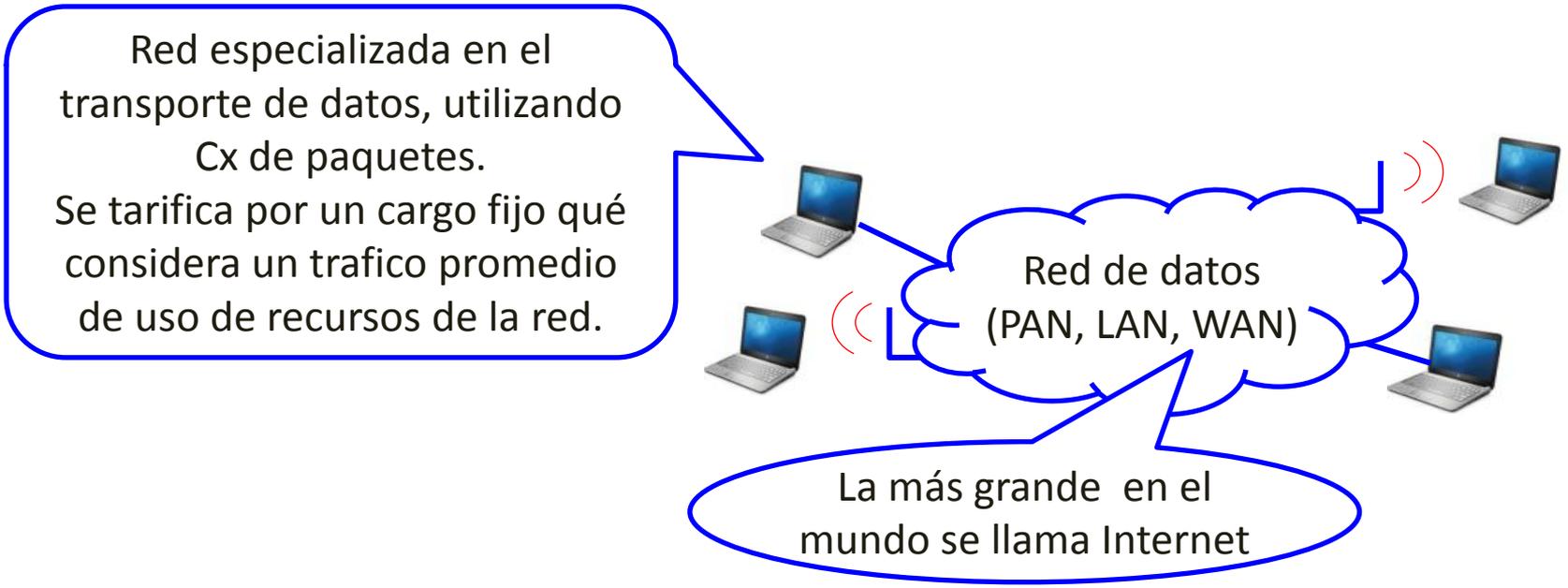
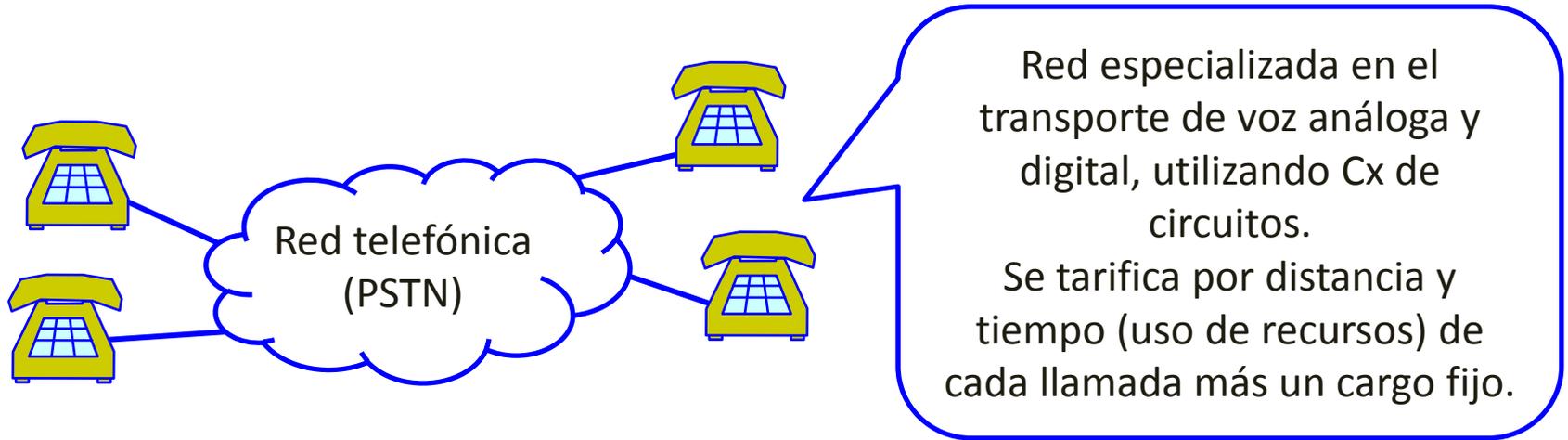
3: Características y requerimientos de una red IP para el transporte de señales de voz, audio y video

4: Estándares y normativa de sistemas de procesamiento y transmisión de voz y video IP

Clasificación del Servicio Telefónico



La naturaleza de las redes telefónica y de datos





Cuando se desea utilizar la PSTN para transmitir datos se encuentra con dos limitaciones:

- La Cx de circuitos genera un costo mayor por uso de la red.
- La velocidad queda limitada a $n \times 64$ Kbps, siendo n el número de canales telefónicos a utilizar.

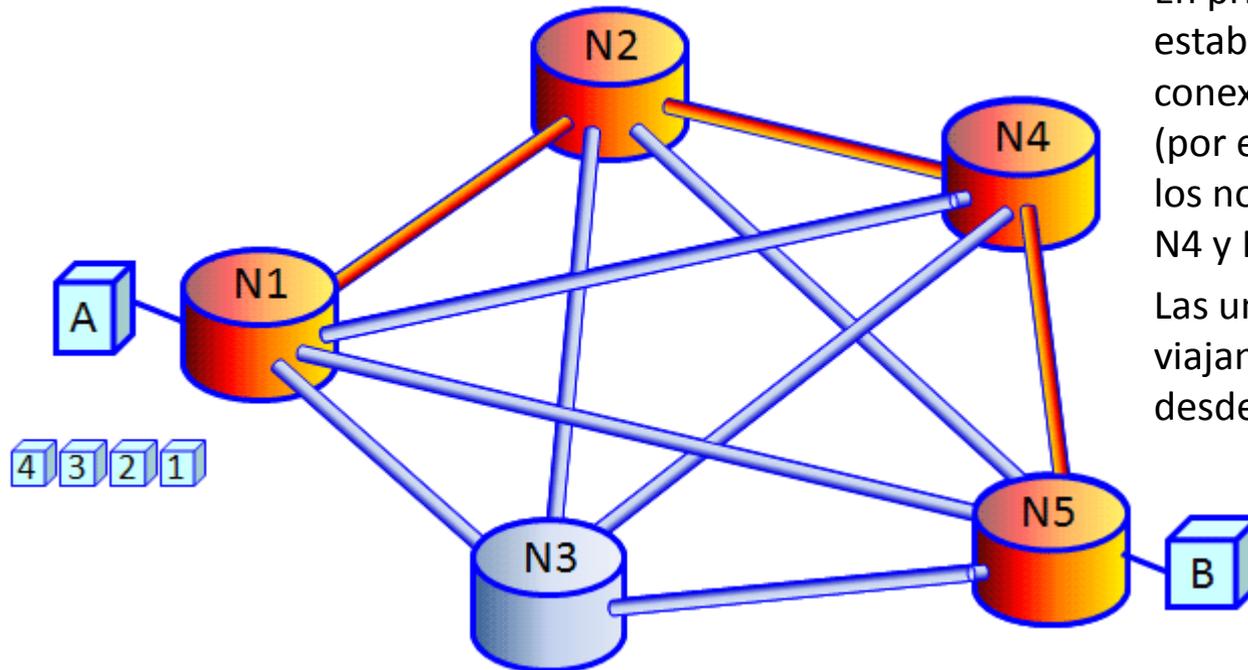
Cuando se desea utilizar la red de datos para transmitir voz encuentra con dos limitaciones:

- La Cx de paquetes no garantiza la secuencialidad de los paquetes y un retardo aceptable en la comunicación.
- La calidad y continuidad de la comunicación es menor que en PSTN.



Una posibilidad de transmitir voz es asegurar las condiciones para emular la Cx de circuitos en una red de Cx de paquetes.

Tipos de conmutación: Conmutación de circuitos



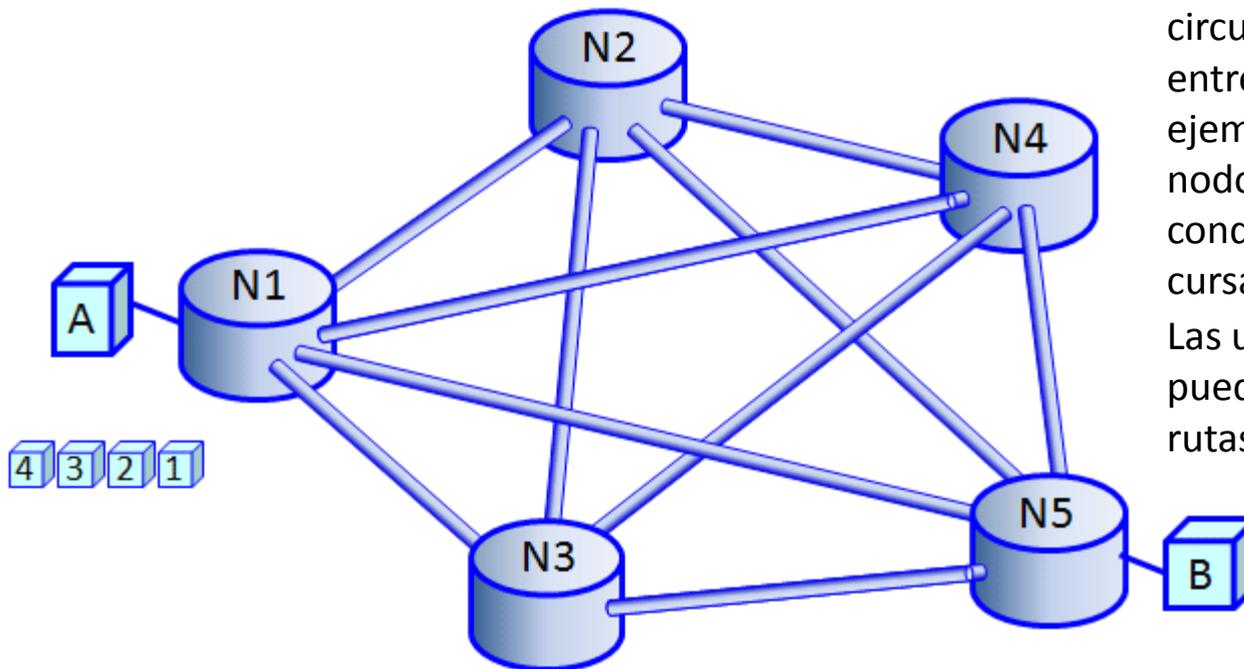
En primer lugar se requiere establecer un circuito físico de conexión entre las estaciones (por ejemplo A y B a través de los nodos o centrales N1, N2, N4 y N5).

Las unidades de información viajan por la ruta preestablecida desde origen a destino.

La ventaja es disponer de un canal de comunicación dedicado, haciéndolo muy adecuado para comunicaciones on line (voz y/o video).

La desventaja es que la reserva de recursos de la red aumenta el costo y no permite aprovechar los periodos de tiempo cuando no se utiliza el canal (momentos de silencio o pausa de la comunicación.)

Tipos de conmutación: Conmutación de paquetes



No se requiere establecer un circuito físico de conexión entre las estaciones (por ejemplo A y B). Todos los nodos deben estar en condiciones y dispuestos para cursar tráfico de información. Las unidades de información pueden viajar por diferentes rutas desde el origen al destino.

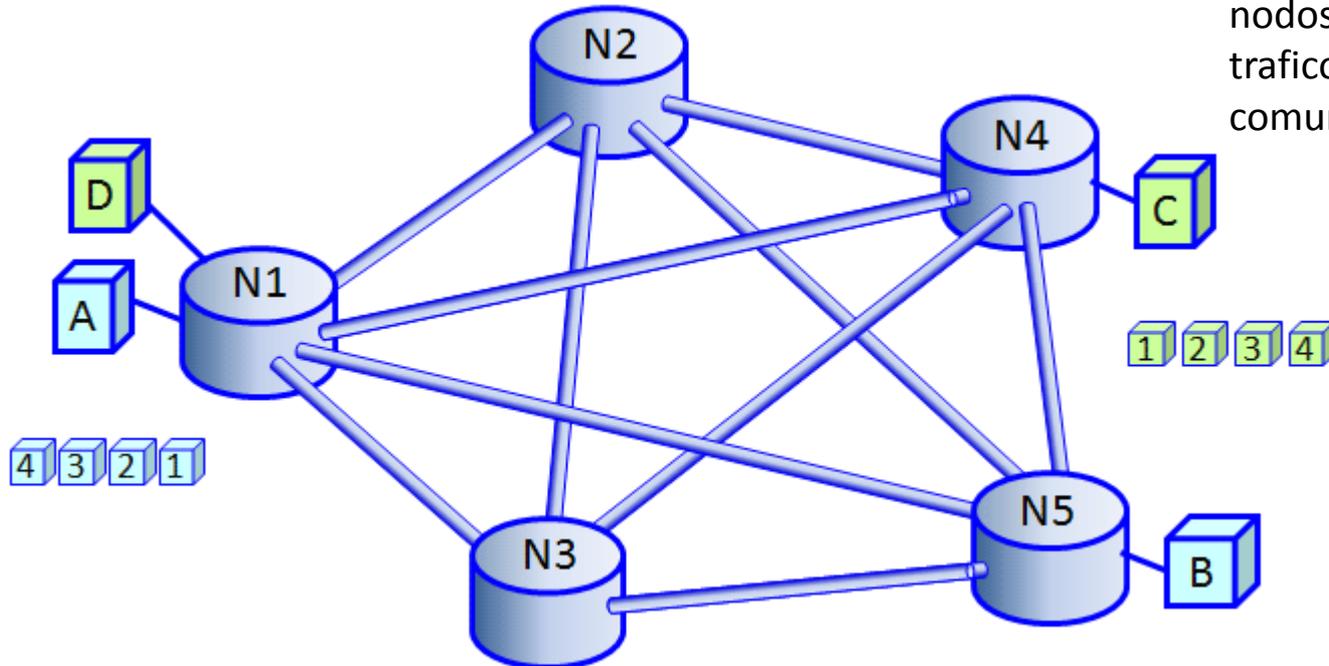
La ventaja es menor costo por tiempo de uso de la red y mayor aprovechamiento de la red. La desventaja es que se pueden producir retardos o desorden de la secuencia de unidades de información enviadas.

En este tipo de conmutación se puede programar para que las unidades de información sigan una ruta similar a la conmutación de circuitos, pero sin reserva o dedicación de nodos y canales a dicha comunicación.

Conmutación de paquetes

Tipos de conmutación: Conmutación de paquetes

La conmutación de paquetes permite que nodos y enlaces cursen tráfico de diferentes comunicaciones.



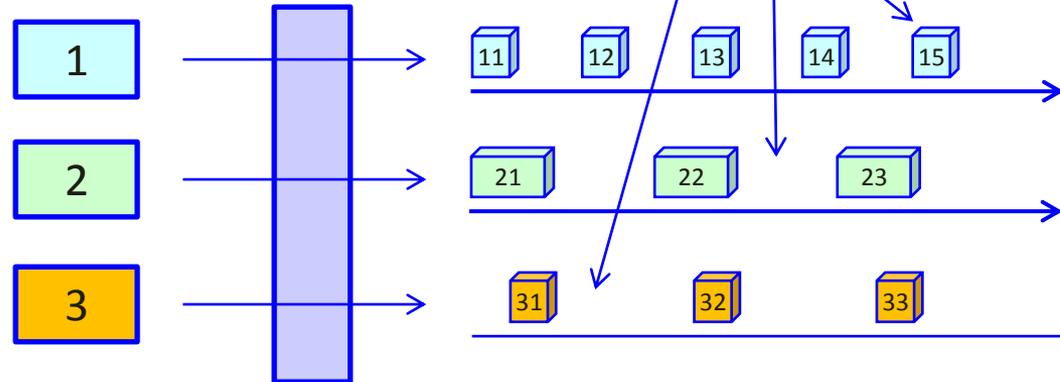
La conmutación de paquetes permite que por un enlace circulen en forma simultánea diferentes comunicaciones aprovechando la capacidad del canal en una forma más eficiente. Este tipo de conmutación es muy adecuado para interconexión de computadores, cuyas unidades de información (paquetes) se transfieren como se ha mostrado en este diagrama.

Cx Circuitos versus Cx de paquetes



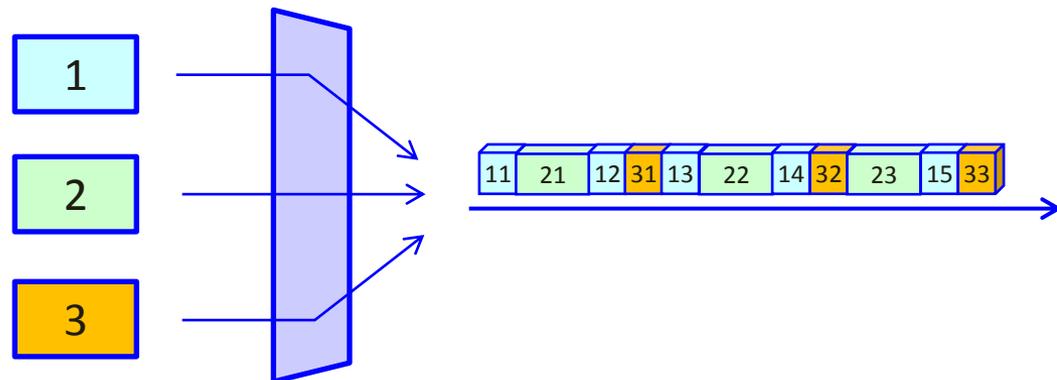
Cx de circuitos:

Cada comunicación utiliza un canal dedicado.
Se generan tiempos muertos en cada canal.
Se produce uso deficiente del BW.

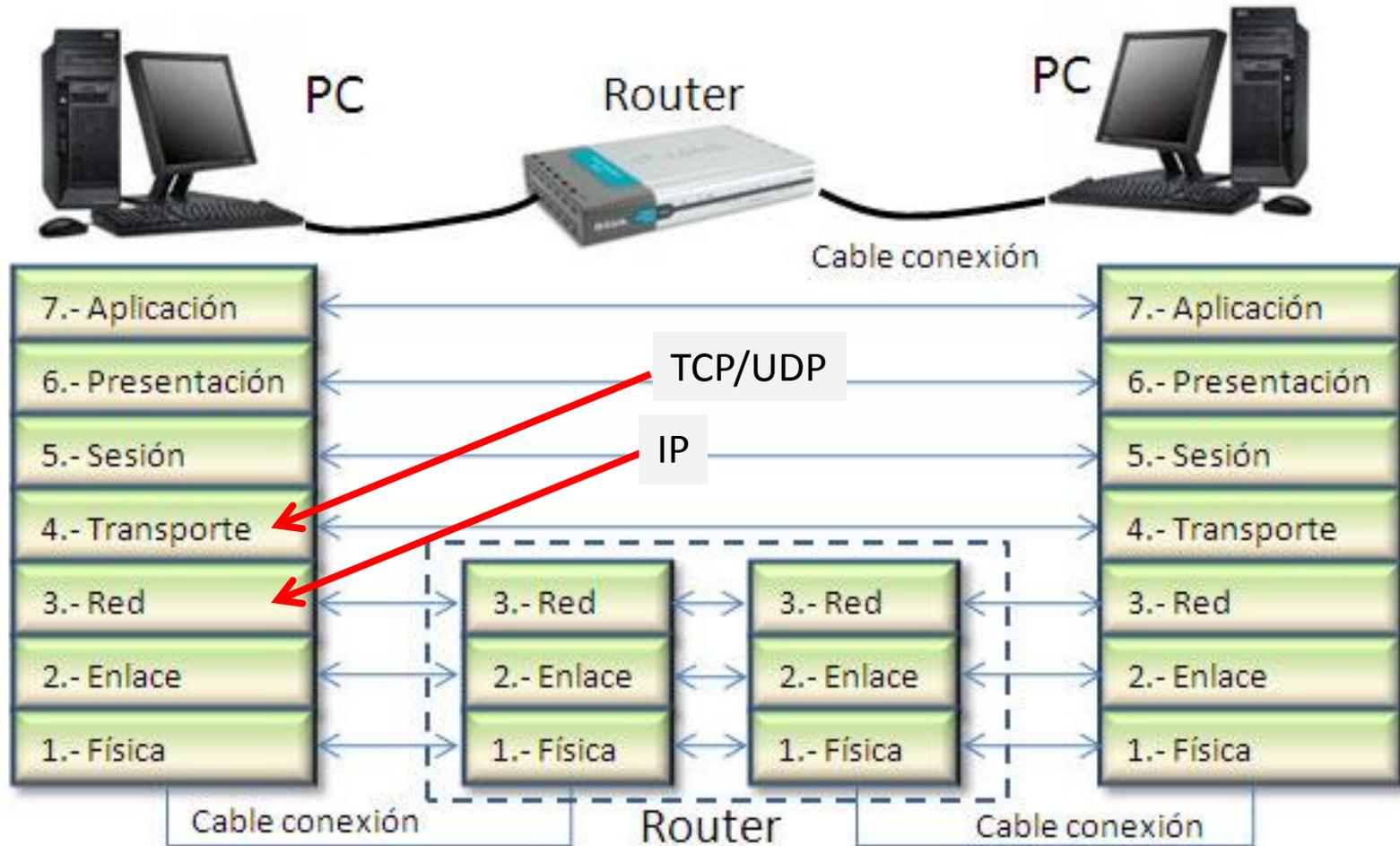


Cx de paquetes:

Las comunicaciones se distribuyen en forma estadística mediante multiplexación asíncrona.
Optimización en el uso del BW.



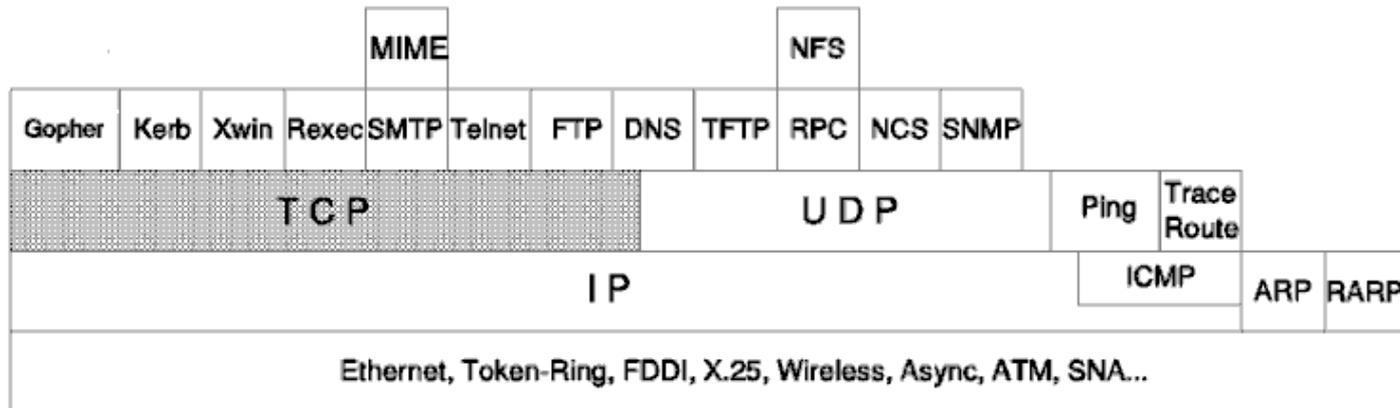
Modelo OSI de una red de datos



Protocolos: TCP



TCP es un protocolo estándar que se describe en el RFC 793 - Protocolo de Control de Transmisión. Este protocolo se recomienda, pero en la práctica cada implementación de TCP/IP que no se use exclusivamente para encaminamiento incluirán TCP.



TCP proporciona muchas más facilidades para las aplicaciones que UDP, recuperación de errores, control de flujo y confiabilidad. Es un protocolo orientado a conexión en contraposición a UDP que no lo es.

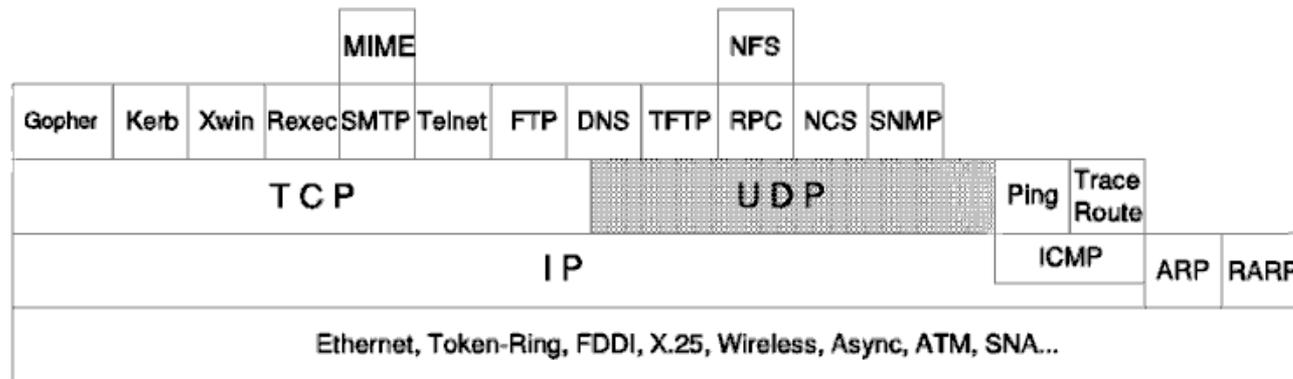
La mayoría de los protocolos de las aplicaciones de usuario, como TELNET y FTP, utiliza TCP.

Protocolos: UDP

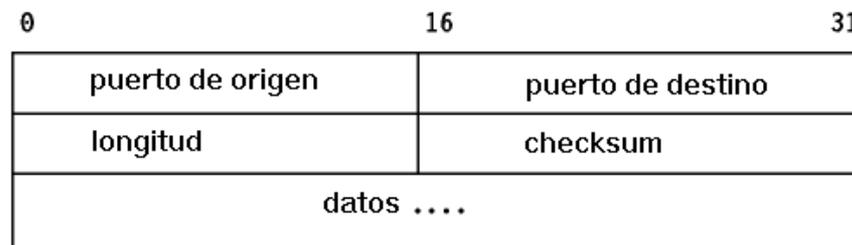


UDP es un protocolo estándar que se describe en el RFC 768 - Protocolo de Datagrama de Usuario.

Este protocolo se recomienda, pero en la práctica cada implementación TCP/IP que no se use exclusivamente para encaminamiento incluirán UDP.



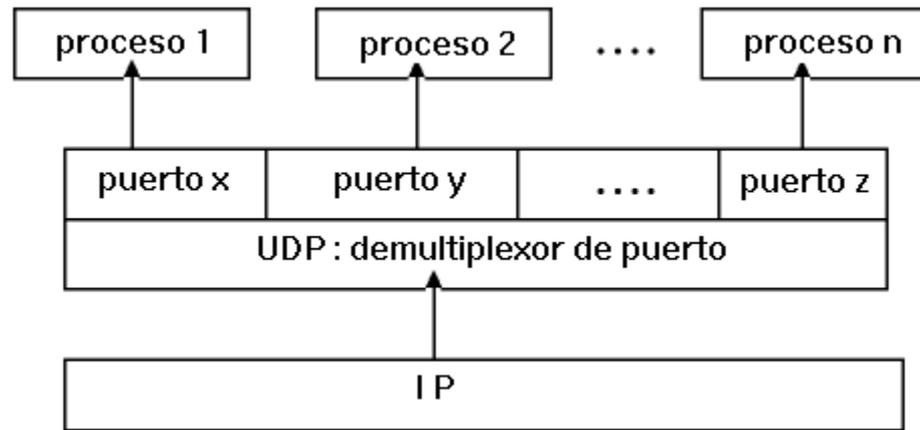
UDP es básicamente una interfaz de aplicación para IP. No soporta confiabilidad, control de flujo o recuperación de errores para IP.



Protocolos: UDP



Simplemente sirve como "multiplexor/demultiplexor" para enviar y recibir datagramas, usando *puertos* para dirigir los datagramas.



UDP proporciona un mecanismo para que una aplicación envíe un datagrama a otra. La capa UDP es sumamente delgada por lo que tiene pocas sobrecargas, pero requiere que la aplicación sea responsable de la recuperación de errores y demás características no soportadas.



Calidad de Servicio (QoS)

- QoS (Quality of Service): Calidad de servicio es cuando un proveedor de servicios garantiza un valor límite (máximo o mínimo) de alguno de los parámetros de QoS.
- Si el proveedor no se compromete en ningún parámetro decimos que lo que ofrece un servicio 'best effort' (mejor esfuerzo).
- El contrato que especifica los valores acordados entre el proveedor y el usuario (cliente) se denomina SLA (Service Level Agreement). Ejemplo:
 - Ancho de banda ≥ 2 Mb/s
 - Retardo ≤ 80 ms
 - Jitter ≤ 20 ms
 - Tasa de pérdidas $\leq 0,01$ %

Parámetros de Calidad de Servicio



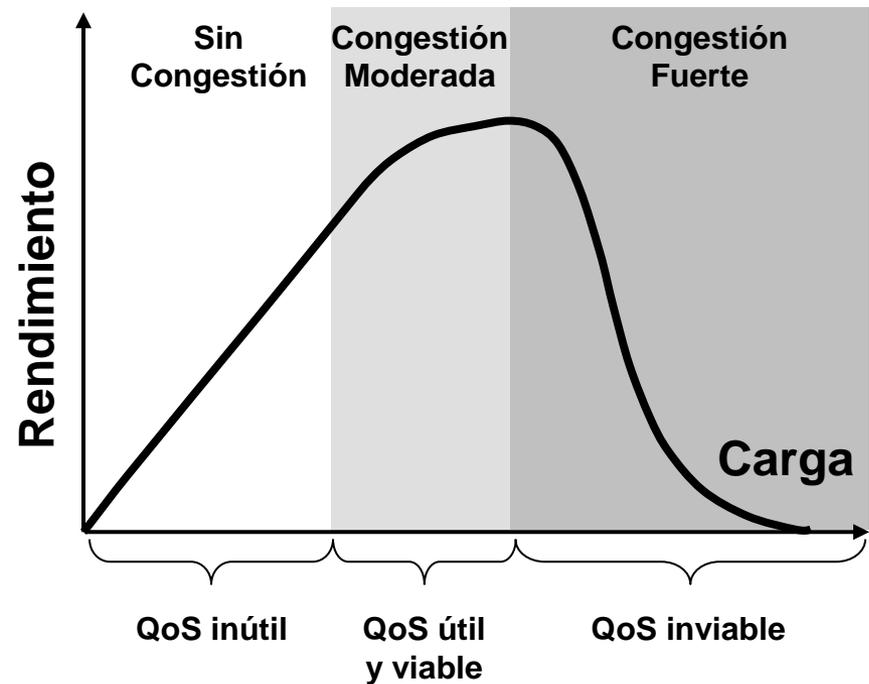
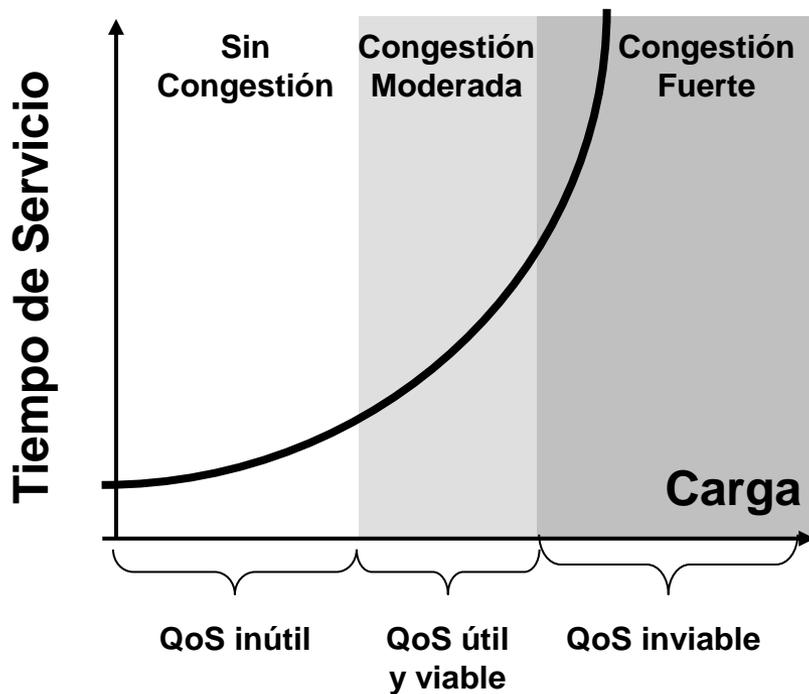
- A continuación se indican los parámetros de afectación más importantes sobre una señal de voz en una red de datos:

Parámetro	Unidades	Significado
Ancho de Banda (bandwidth)	Kb/s	Indica el caudal máximo que se puede transmitir
Retardo (delay) o latencia (latency)	ms	El tiempo medio que tardan en llegar los paquetes
Jitter	ms	La fluctuación que se puede producir en el Retardo
Tasa de pérdidas (loss rate)	%	Proporción de paquetes perdidos respecto de los enviados

Efectos de la congestión

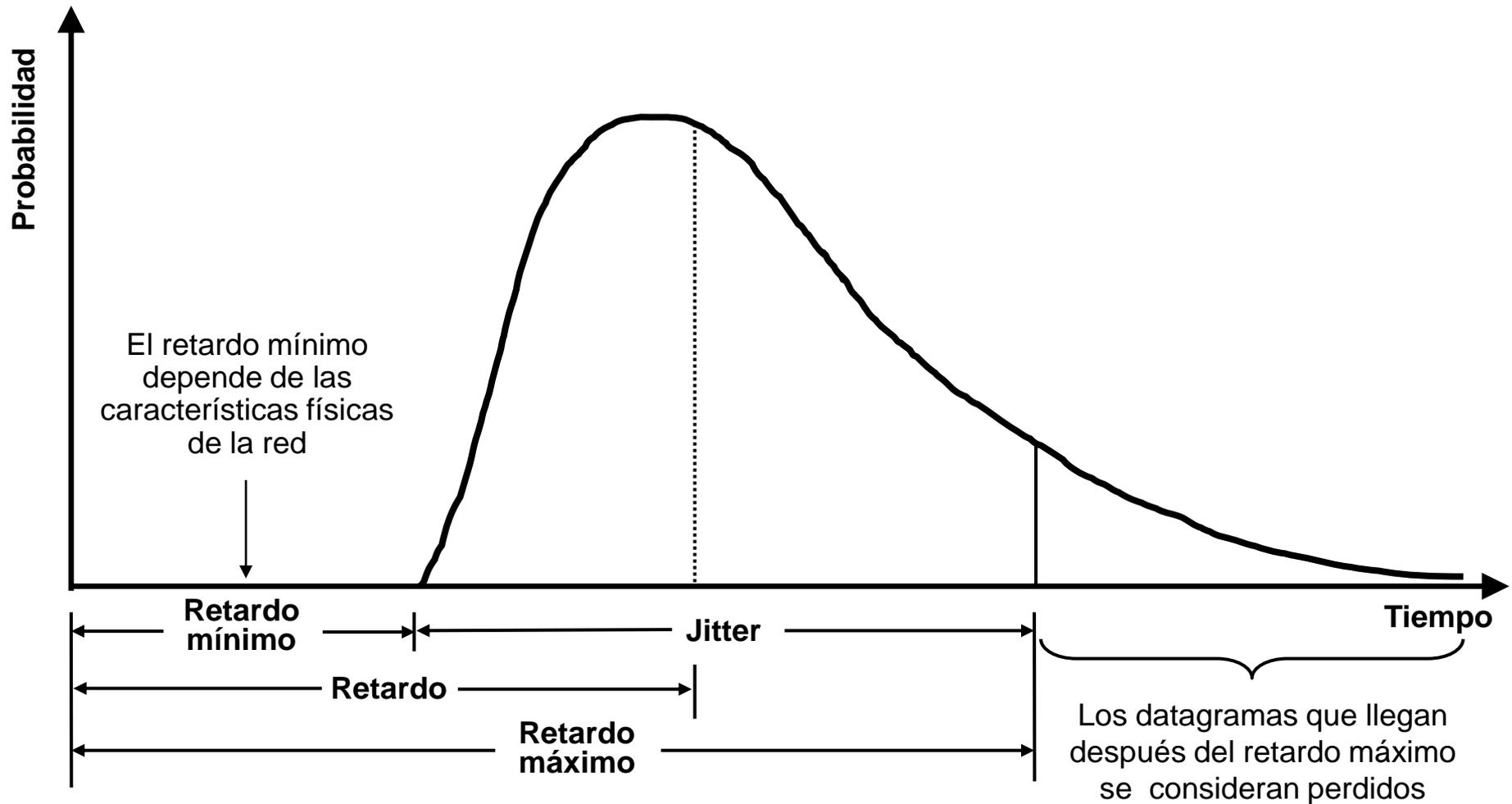


- Para dar QoS con congestión es preciso tener mecanismos que permitan dar un trato distinto al tráfico preferente y cumplir el SLA (Service Level Agreement).



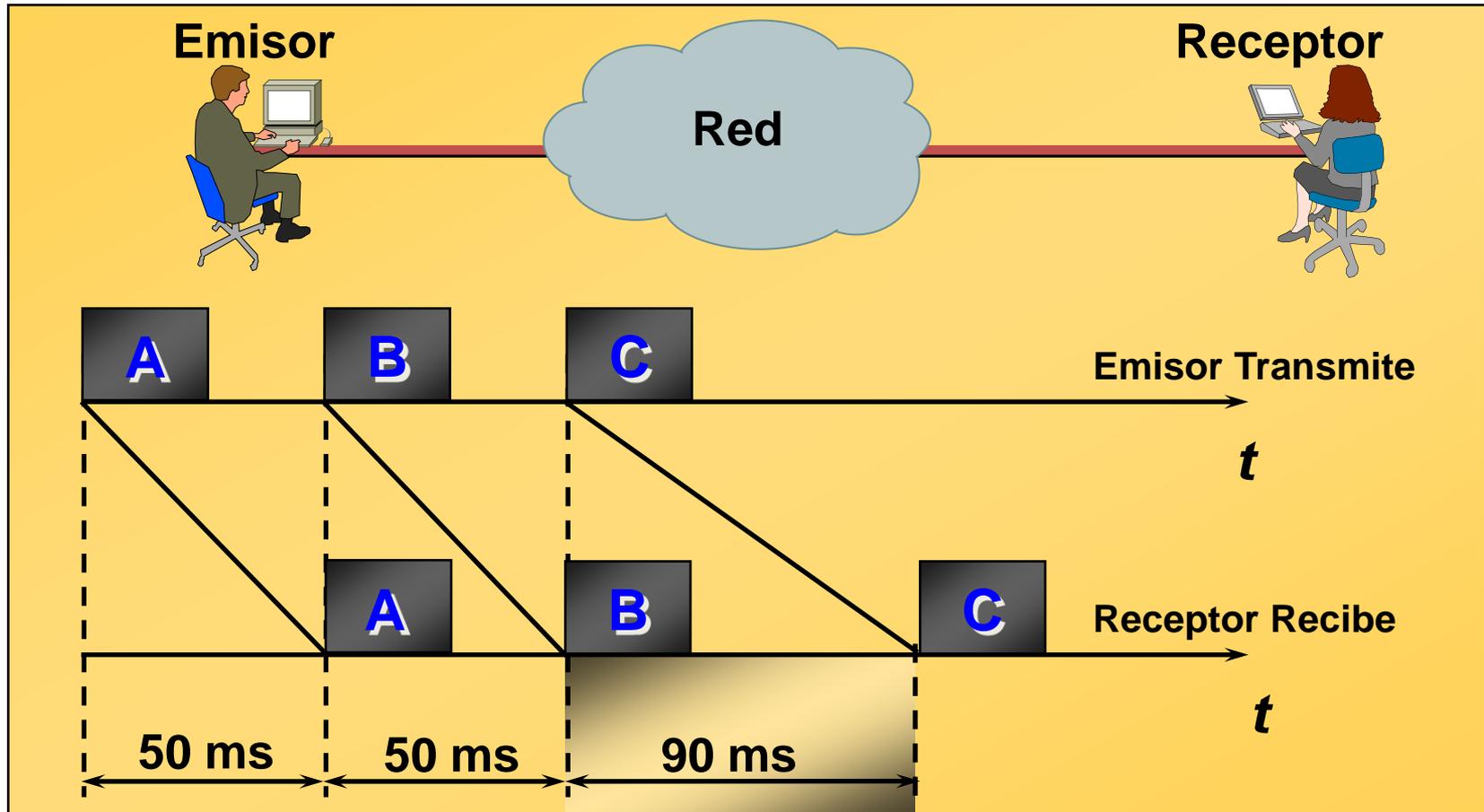
Fuente: http://www.uv.es/montanan/ampliacion/amplif_6.ppt

Relación entre la probabilidad de llegada de los paquetes y los parámetros de QoS



Fuente: http://www.uv.es/montanan/ampliacion/amplif_6.ppt

Concepto de Jitter



Red vacía Congestión

Retardo: $70 \text{ ms} \pm 20 \text{ ms}$ (retardo: 70 ms, jitter: 40 ms)

Fuente: http://www.uv.es/montanan/ampliacion/amplif_6.ppt



- El jitter puede reducirse si el receptor retrasa la reproducción (buffer 'anti-jitter').
- Por ejemplo en VoIP lo habitual es enviar un paquete de voz cada 20 ms. Si el receptor reproduce los paquetes tal cual le llegan cualquier fluctuación en la entrega afectará la calidad. Si en vez de eso retrasa 40 ms la reproducción podrá compensar fluctuaciones de hasta 40 ms en el tiempo de entrega.
- En algunas aplicaciones (vídeo o audio unidireccional) se llegan a introducir retardos de hasta 30 segundos. Pero en estos casos no existe interacción receptor-emisor

Tipo de aplicación	Ancho de Banda	Retardo	Jitter	Tasa de Pérdidas
Interactivo (telnet, www)	Bajo	Bajo	Medio/alto	Media ¹
Batch (e-mail, ftp)	Alto	Alto	Alto	Alta ¹
Telefonía	Bajo	Bajo	Bajo	Baja
Vídeo interactivo	Alto	Bajo	Bajo	Baja
Vídeo unidireccional (streaming)	Alto	Medio/alto	Bajo	Baja
Frágil (ej.: emulación de circuitos)	Bajo	Bajo	Medio/alto	Nula

¹En realidad la aplicación requiere pérdida nula, pero esto lo garantiza el protocolo de transporte TCP

De este modo se presentan dos alternativas para ofrecer calidad de servicio: Reserva de recursos de red o prioridad en el flujo de tráfico de la red.

¿Reserva o Prioridad?



	Ventajas	Inconvenientes
Reserva	<ul style="list-style-type: none">•Da una garantía casi total•Los paquetes no necesitan llevar ninguna marca que indique como han de ser tratados, la información la tienen los routers	<ul style="list-style-type: none">•Requiere mantener información de estado sobre cada comunicación en todos los routers por lo que pasa•Se requiere un protocolo de señalización para informar a los routers y efectuar la reserva en todo el trayecto
Prioridad	<ul style="list-style-type: none">•Los routers no necesitan conservar información de estado.	<ul style="list-style-type: none">•Los paquetes han de ir marcados con la prioridad que les corresponde•La garantía se basa en factores estadísticos, es menos segura que la reserva de recursos (puede haber overbooking)

Fuente: http://www.uv.es/montanan/ampliacion/amplif_6.ppt

Calidad de Servicio en Internet



- La congestión y la falta de QoS es el principal problema de Internet actualmente.
- IP fue diseñado para dar un servicio 'best effort'. Sin embargo hoy en día se utiliza para aplicaciones sensibles que no toleran redes sin QoS. Ej.: videoconferencia, telefonía VoIP (Voice Over IP), etc.
- Estas aplicaciones no pueden funcionar en una red 'best effort' congestionada.
- Se han hecho modificaciones a IP para que pueda ofrecer QoS a las aplicaciones mencionadas.



Siglas de Real-time Transfer Protocol

RTP es ampliamente usado para transportar flujos de datos correspondiente a audio y video.

Está diseñado para aplicaciones que envían datos en una dirección sin señal de reconocimiento (sin acknowledgement)

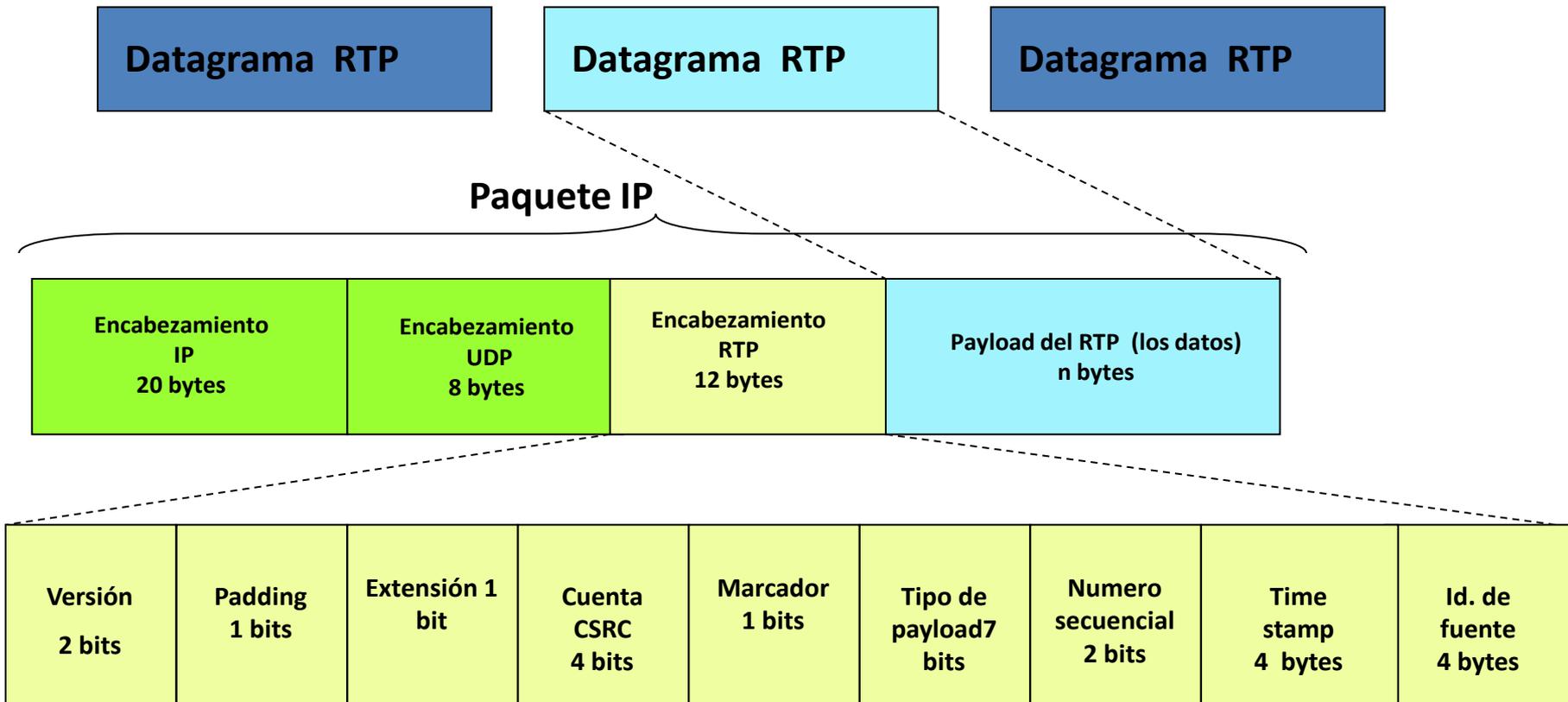
El encabezamiento de cada datagrama RTP contiene un “timbre con la hora” (timestamp). Este “timestamp” permite que la aplicación en el extremo receptor reconstruya exactamente el “timing” de la información original

El encabezamiento también contiene un número secuencial, que permite al extremo receptor detectar y tomar medidas cuando los datagramas sufren pérdida, duplicación o deterioro

Protocolo RTP



En el paquete IP el encabezamiento de RTP va a continuación del encabezamiento de UDP y éste a continuación del encabezamiento de IP





Si bien se habla en general de Voz sobre IP (VoIP: Voice over Internet Protocol), al mencionar la tecnología debemos hacer las salvedades del caso y distinguir dentro de la voz paquetizada, dos aplicaciones con mercados y características distintas.

Voz sobre IP (VoIP): se emplea dicho término a toda implementación de voz paquetizada que se utiliza en una red privada (LAN) pudiendo esta tener o no contacto con la PSTN.

Telefonía sobre IP (ToIP): se emplea dicho término a toda red de voz paquetizada, soportada sobre redes de área amplia (WAN), las cuales cumplen las veces de la PSTN en forma total y/o parcial. En general la tendencia muestra que las redes de ToIP se tratan como redes de servicio público y por lo tanto se deben regir por normas y estándares establecidos por las entidades administradoras de telecomunicaciones.

1: Principios físicos de imagen y sonido y fisiológicos de audición y visión humana y su conversión para el procesamiento y transmisión

2: Procesamiento y transmisión analógico y digital de señales de voz, audio, imagen y video

3: Características y requerimientos de una red IP para el transporte de señales de voz, audio y video

4: Estándares y normativa de sistemas de procesamiento y transmisión de voz y video
IP

Los principales organismos de estandarización de VoIP, son:

ITU



IETF



IMTC



ETSI



- La ITU presenta en el año 1996, la especificación H.323. Se dice que H.323 es una especificación ya que corresponde a conjunto de normas agrupadas.
- La razón de esto es que la H.323, no solamente cubre lo relacionado con la voz, sino también con lo relativo a protocolos y “Multimedia”, debiendo especificar entonces el tratamiento de video, datos su multiplexado y control.
- Por otra parte, IETF desarrollo en paralelo el protocolo SIP (Session Initiation Protocol) orientado a conexiones multimedia.

Los estándares de video y televisión analógica



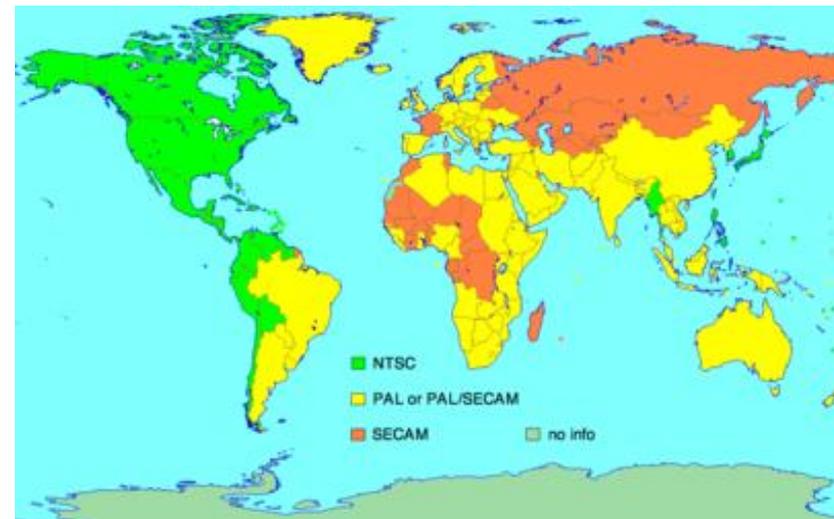
En el mundo existen actualmente 3 sistemas de televisión analógica:

NTSC (National Television System Committee) Sistema de codificación y transmisión de televisión a color analógica desarrollado desde 1940. Se emplea en la actualidad en la mayor parte de América y Japón, entre otros países.

PAL (Phase Alternating Line). Sistema de codificación y transmisión de señales de televisión analógica en color. Se utiliza en la mayoría de los países africanos, asiáticos y europeos, además de Australia y algunos países latinoamericanos siendo el de mayor cobertura y uso en el mundo.

SECAM (Séquentiel Couleur avec Mémoire) "Color secuencial con memoria". Es un sistema para la codificación de televisión en color analógica utilizado en Francia, algunos países de África y Asia.

En muchos países del mundo se esta migrando a la televisión digital.

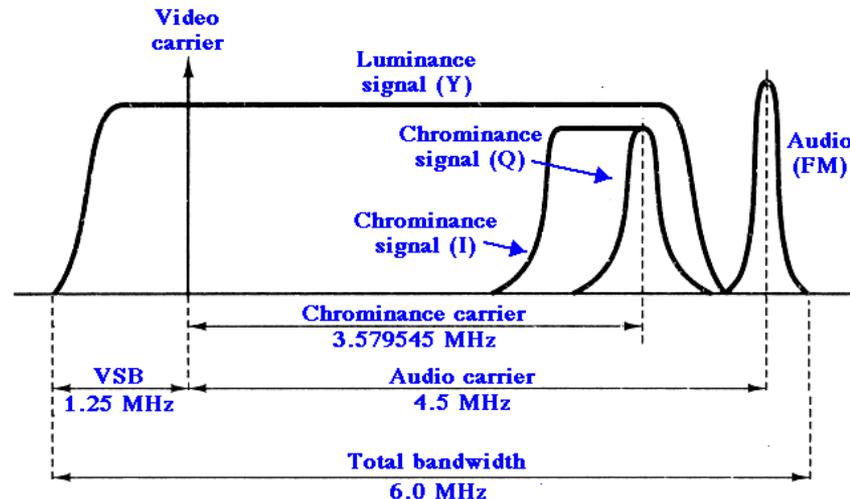


Estándar NTSC



La actual norma de Televisión NTSC fue desarrollado en 1953 y considera la transmisión de una señal compuesta por 3 portadoras moduladas distribuida en un BW de 6 MHz, como se indica:

- Imagen de video a razón de 30 cuadros por segundo (30 fps)
- Video modulada en amplitud que contiene la información de luminancia Y.
- Croma o color modulada en cuadratura de fase (I, Q) que contiene en forma combinada la información de los colores básicos (RGB).
- Audio modulada en frecuencia para dos canales (estéreo).
- Barrido entrelazado.

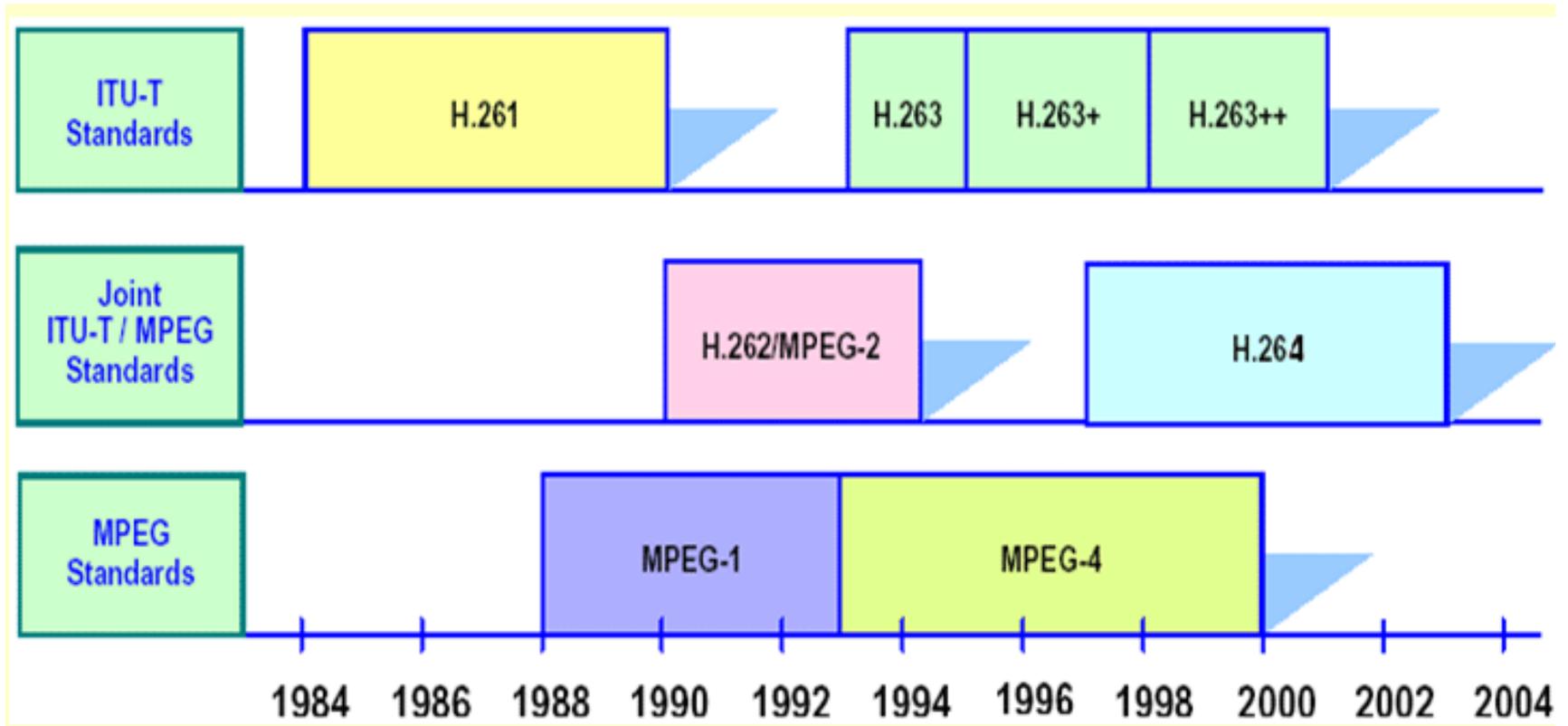


También establece la relación de ancho - alto de la pantalla en 4:3 para hacerla compatible con el formato del cine de la época.

Resolución de video para NTSC

- ✓ La cantidad de líneas de barrido estándar NTSC es de 525.
- ✓ En el caso de las imágenes de vídeo NTSC, el estándar profesional SMPTE 259M especifica que las 525 líneas se representa como 720 x 486: es decir, 720 píxeles horizontales por 486 píxeles verticales. El tamaño de vídeo predeterminado se conoce comúnmente como D1.
- ✓ La captura de imágenes mediante una fuente DV (vídeo digital), sin embargo, proporciona un fotograma de 720 x 480. La diferencia entre la especificación D1 y la especificación DV es de tan sólo 6 píxeles verticales. Muchos algoritmos de compresión, incluida la compresión DV, funcionan mejor cuando los tamaños de imagen son múltiplos de 16. Al quitar los 6 píxeles de una resolución D1, el formato DV se puede disponer de una resolución nativa con un múltiplo de 16.

Evolución de estándares de video



Formatos y estándares de video digital



Formato	Compres. Video	Resolución PAL	Resolución NTSC	Caudal Video (Mbps)	Compres. Audio	Caudal Audio (Kbps)	Cons. CPU	Calidad
ASF, SMR/nAVI	MPEG-4	320 x 240 o menor	320 x 240 o menor	0,1-0,5	MPEG-4	64-128	Alto	Mala
VCD	MPEG-1	352x288	352 x 240	1,15	MPEG-1 (II)	224	Bajo	Regular
XVCD	MPEG-1	352x288	352 x 240	1,5-2,5	MPEG-1 (II)	32-384	Bajo	Regular
CVD	MPEG-2	352x576	352 x 480	1,5-2,5	MPEG-1 (II)	128-384	Medio	Buena
SVCD	MPEG-2	480 x 576	480 x 480	1,5-2,5	MPEG-1 (II)	32-384	Medio	Buena
DivX	MPEG-4	640 x 480 o menor	640 x 480 o menor	0,3-1	MPEG-1 (III), WMA	64-192	Muy alto	Buena
XSVCD	MPEG-2	352 x 288 352 x 576 720 x 576	352 x 240 352 x 480 720 x 480	1,5-3	MPEG-1 (II)	32-384	Medio	Buena
DVD	MPEG-2	720 x 576	720 x 480	3-8	MPEG-2	192-448	Alto	Excelente
DV	DV	720 x 576	720 x 480	25	DV	1000-1500	Alto	Excelente

- ✓ Estándares de vídeo la ITU-T para videoconferencia: baja velocidad, poco movimiento. Menos acción que en el cine.
 - H.261: Desarrollado a finales de los 80 para RDSI (caudal constante).
 - H.263, H.264, H.265. Más modernos y eficientes.
- ✓ Algoritmos de compresión MPEG simplificados:
 - Vectores de movimiento más restringidos (menos acción)
 - En H.261: No fotogramas B (excesiva latencia y complejidad)
- ✓ Menos intensivo de CPU. Factible codec software en tiempo real
- ✓ Submuestreo 4:1:1
- ✓ Resoluciones:
 - CIF (Common Interchange Format): 352 x 288
 - QCIF (Quarter CIF): 176 x 144
 - SCIF (Super CIF): 704 x 576
- ✓ Audio independiente: G.722 (calidad), G.723.1, G.728, G.729
- ✓ Sincronización audio-vídeo mediante H.320 (RDSI) y H.323 (Internet)

Estándar H264 de ITU



Aplicaciones:

CATV Televisión por cable en redes ópticas, de cobre, etc.
(*cable TV on optical networks, copper, etc.*)

DBS Servicios de vídeo de radiodifusión directa por satélite
(*direct broadcast satellite video services*)

DSL Servicios de vídeo por línea de abonado digital (*digital subscriber line video services*)

DTTB Radiodifusión de televisión terrenal digital (*digital terrestrial television broadcasting*)

ISM Medios de almacenamiento interactivos (discos ópticos)
(*interactive storage media (optical disks, etc.)*)

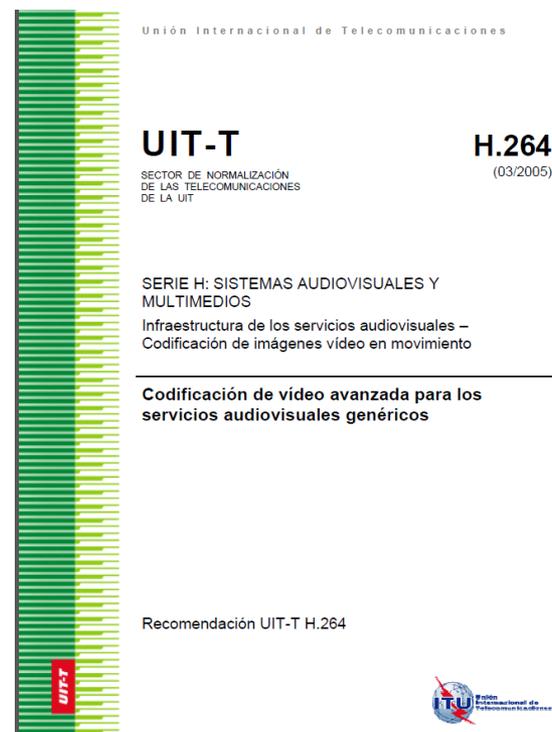
MMM Distribución de correo multimedia (*multimedia mailing*)

MSPN Servicios multimedia por redes de paquetes (*multimedia services over packet networks*)

RTC Servicios conversacionales en tiempo real (videoconferencia, videoteléfono, etc.)
(*real-time conversational services (videoconferencing, videophone, etc.)*)

RVS Televigilancia por vídeo a distancia (*remote video surveillance*)

SSM Medios de almacenamiento en serie (VTR digital, etc.) (*serial storage media (digital VTR, etc.)*)





Esta Recomendación (Norma Internacional) es una evolución de las normas de codificación de vídeo existentes (H.261, H.262 y H.263) y se elaboró para cubrir la creciente necesidad de mayor compresión de imágenes en movimiento en diversas aplicaciones como videoconferencia, medios de almacenamiento digital, radiodifusión de televisión, emisión por Internet y comunicaciones.

Asimismo, se ha concebido para que la representación de vídeo codificado se pueda utilizar de manera flexible en una gran variedad de entornos de red. La utilización de esta Recomendación permite manipular imágenes de vídeo como simples datos digitales, almacenarlas en diversos tipos de medios de almacenamiento, transmitir las y recibir las por las redes existentes y futuras, y distribuir las por los canales de radiodifusión existentes y futuros.

La revisión aprobada en 2005-03 contiene modificaciones a la norma de codificación de vídeo para añadir cuatro nuevos perfiles, denominados alto, alto 10, alto 4:2:2 y alto 4:4:4, a fin de mejorar la capacidad de calidad de vídeo y ampliar una gama de aplicaciones tratadas por la norma (por ejemplo, incluyendo el soporte de una mayor gama de precisión de muestras de imágenes y de formatos croma de más alta resolución).

Definiciones Subtel de VoIP y ToIP:

¿Qué es el Servicio Público de Voz sobre Internet?

Es un servicio de telecomunicaciones que permite la prestación de comunicaciones de voz sobre la red Internet desde y hacia la red pública telefónica u otra red de servicio público del mismo tipo.

¿Qué es la telefonía IP?

La telefonía IP (Internet Protocol) es una tecnología que reúne la transmisión de voz y de datos, posibilitando la utilización de redes informáticas para efectuar llamadas telefónicas. Desarrolla una red única, que se encarga de cursar todo tipo de comunicación, ya sea, de voz, datos o video, que se denomina Red Convergente o Red Multiservicios.

La telefonía IP surge como una alternativa a la telefonía tradicional, brindando nuevos servicios al cliente entre los que se cuentan una serie de beneficios económicos y tecnológicos.

Respecto de la telefonía IP, Subtel señala que la regulación indicada en el [Decreto 484/2007](#) sólo se refiere a aquellos servicios de voz que se prestan sobre Internet y que cumplen con las siguientes condiciones:

- que realicen llamados a la Red Pública Telefónica y
- que reciban llamados desde la Red Pública Telefónica.

En este caso, el proveedor de servicios requiere una concesión.

Por lo tanto, este reglamento no regula las llamadas entre usuarios de Internet, por ejemplo las que se efectúan vía MSN u otras prestaciones similares y tampoco regula servicios que permiten hacer llamados a la Red Pública, pero sin posibilidad de recibir, como ocurre con Skype out y equivalentes.

Este marco legal, por tanto, permitirá a los usuarios de Internet acceder a servicios de comunicación desde y hacia la red pública telefónica, teniendo el derecho a reclamar por mal servicio, a recibir una cuenta detallada, a llamar en todo momento a números de emergencia -aún cuando el servicio esté suspendido por no pago- y a poner término al contrato de suministro en forma unilateral en un plazo máximo de 10 días, entre otros derechos.

Derechos y obligaciones del concesionario de Servicio Público de VoIP:

1. El servicio de voz sobre Internet se considera un Servicio Público y su prestación requiere una concesión.
2. Se establece un bloque de numeración con características ageográficas, es decir, sin importar donde el usuario esté conectado a Internet, mantiene su número (similar a la telefonía móvil).
3. El concesionario deberá informar al usuario respecto a la calidad de servicio que está prestando.
4. El servicio no está sujeto al uso de multiportador para comunicaciones LD.
5. No tiene guía telefónica
6. Podrán interconectarse con la red pública telefónica directa o indirectamente usando medios propios o de terceros.
7. Está sujeto a la interceptación telefónica conforme a las normativas de seguridad pública establecidas por ley.
8. La implementación del acceso a los servicios de emergencia es obligatoria y permanente.
9. Las responsabilidades del concesionario del servicio, se separan de las del prestador del servicio de banda ancha y del proveedor del acceso Internet (ISP).
10. Están sujetos a cumplir el Reglamento de Resolución de Reclamos.
11. Están sujetos a la obligación de entregar una Cuenta Única.