7.1 Introduccion a Multimedia

Multimedia: Información con más de un medio. Se asociará la definición de multimedia a aquellos datos o archivos que tengan más de un medio y sean reproducibles por un intervalo de tiempo.

Película: información multimedia formada por audio y video (imágenes en movimiento), en señales diferentes.

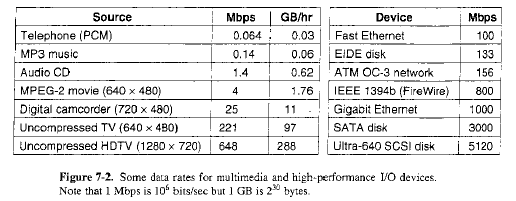
Ejemplos de medios de almacenamiento:

DVD: 4.5 a 17Gb.  
Blue Ray: 25Gb en single layer y 50Gb en double layer.  
HD DVD: 15Gb en single Layer y 30Gb en double layer.

2 caracteristicas fundamentales sobre multimedia:

* Multimedia utiliza muy altas velocidades de transmisión de datos
* Multimedia requiere reproducción en tiempo real.

Formatos multimedia con sus tamaños / transferencias.



NTSC 🡪 29.97 fps  
PAL – SECAM 🡪 25 fps

Se debe respetar el tiempo entre frames o se notara en la pelicula los saltos. (1 frame por cada 33.3 msec y 40msec respectivamente)

“jitter” es el termino usado para cuando hay variabilidad en la entrega del medio (saltos de diferentes tiempos en el recurso multimedia).

Ejemplo de disponibilidad de servicio:

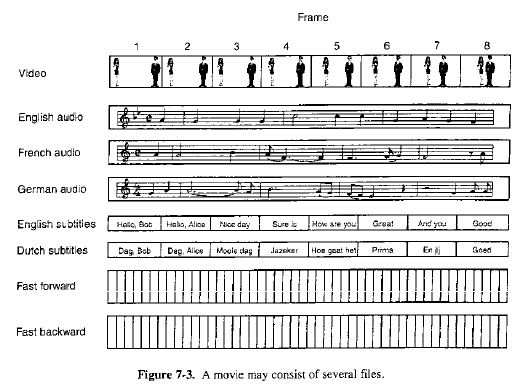
* Ancho de banda: 4Mbps
* 99% de la demora de transmision: 105 a 110ms
* Rango de perdida de bits: 10xe-10

El operador podria dar diferentes rangos de servicios

7.2 Archivos multimedia

Las películas actualmente pueden emitirse en diferentes idiomas incluyendo subtitulos, para escuchar el idioma original de la película con los subtítulos locales y viceversa. Los mismos pueden ser guardados en archivos diferentes de la película.

DVDs por ejemplo pueden almacenar hasta 32 idiomas en audio y subtitulos.



Para la utilización de distintos archivos se debe guardar un tracking en cada uno.

7.2.1. Video Encoding

Los sistemas convencionales muestran alrededor de 25 cuadros por segundo.  
Los televisores blanco y negro barren la pantalla bidimensional como si fuera de una sola dimensión barriendo secuencialmente las líneas y poniendo la intensidad de luz que corresponde.

El barrido de todas las líneas de la pantalla se lo llama “frame” (cuadro). NTSC barre a 29.97 cuadros por segundo mientras que PAL y SECAM lo hacen a 25.  
Si bien 25 es la cantidad adecuada para hacer una película con animación suave, muchas veces esto produce lo denominado “flicker” (parpadeo). Esto es porque en ciertas personas la imagen en la retina desaparece antes que aparezca la nueva).

Para solucionar esto, en vez de duplicar la cantidad de cuadros se optó por dividir el barrido en líneas pares y luego impares. Cada barrido de la mitad de un frame se lo llama “field” (campo) barriendo un total de 50 campos por segundo.

A esta técnica se lo llama entrelazado (“interlacing”).

Cuando no utiliza la técnica de entrelazado se llama “progressive scan” (escaneo progresivo).

Las pantallas a color funcionan igual que las b/n a diferencia que dividen la señal en 3 haces de luces: rojo, verde y azul (RGB). Para transmisión debe usarse esta triple señal como una sola señal compuesta.

Para que las señales en color puedan verse en dispositivos b/n se combina el RGB en otro tipo de señal: 1 de luminosidad (brillo) (“luminance”) y otras 2 de crominancia (color) (“crhominance”).

El ojo humano es mucho mas sensible a la luminosidad que al color, éste ultimo no tiene que ser tan preciso, por tal motivo la luminosidad se puede seguir transmitiendo de la misma frecuencia que en la antigua señal blanco y negro.

Pasando al video digital, el RGB puede mostrarse utilizando 8 bits por cada color (24 bits en total) por cada punto de la pantalla (pixel) dando un total de 16 millones de colores posibles por cada punto.

Desde que los monitores de PC guardan las imágenes en memoria y escanean a velocidades de 75hz (veces o ciclos por segundo) no es necesario utilizar el entrelazado y por tal motivo todos realizan escaneo progresivo ya que tiene los valores suficientes para eliminar el flicker.

**7.2.2. Audio encoding**

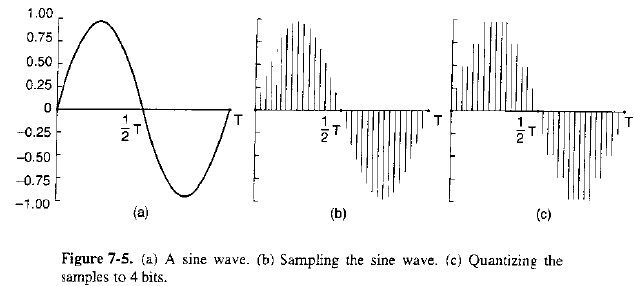
Una onda de sonido es una onda acústica (presión) de una sola dimensión.

El rango de frecuencias del oído humano van desde los 20hz hasta los 20.000hz.

El radio de dos sonidos con amplitudes A y B se expresa en dB (decibeles) según la siguiente formula:



Las ondas de sonido son convertidas en señales digitales por ADC (Analog Digital Converter).



Para representarlo digitalmente cada muestra de sonido lo debe hacer en un intervalo de tiempo t.

Luego cada valor se muestra en un rango finito de valores, en el ejemplo el rango va de -1 a 1 en pasos de 0,25. Para reproducirlos basta con 4 bits. Con 8 bits se pueden representar 256 valores distintos y con 16, 65536.  
El error introducido por este rango finito de valores se llama cuantización de ruido (“quantization noise”). Si este valor es muy alto entonces podría ser captado por el oído.

Ejemplos:

* CD Audio: 44,1KHz 16 bit stereo

**7.3. Compresion de video**

Se necesitan 2 metodos, uno para compresión y otro para descompresión, que en el caso de multimedia pasan a llamarse, “encoder” y “decoder” (codificador, decodificador).

Para la compresión multmedia se permiten asincronismo, es decir, es permisible un tiempo mas alto para la compresión (y con costo mayor de hardware) ya que se realiza por única vez, luego se requiere que la descompresión sea mas rápida y que no tenga altos requerimientos de hard.  
O para videoconferencias, donde se requiere alta velocidad de compresión.

En el caso de la compresión de videos, a diferencia de la compresión normal, es permisible tener cierta perdida de datos, a estos métodos de compresión se los denomina “lossy” ya que pierden algo de información pero comprimen mas.

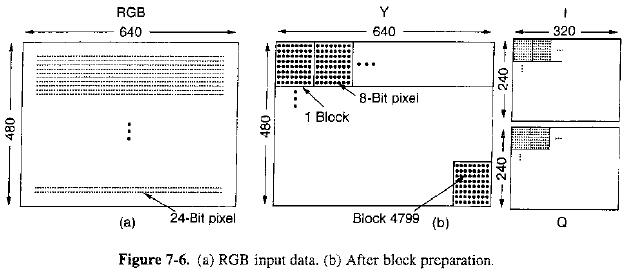
**7.3.1. Standard JPEG**

JPEG = Joint Photographic Expert Group.

JPEG tiene 4 modos pero solo se verá la usada para Video en 24-bits RGB.

Paso 1:

Para codificar en JPEG primero hay que hacer la preparación en bloques. Asumimos una imagen de 640x480x24.  
Se usará luminosidad y crominancia debido a que son mejores para comprimir. Luminosidad y dos canales de crominancia serán tomados desde los valores del RGB (para NTSC los canales son Y, I y q y para PAL Y, U y V). Las fórmulas son diferentes para cada uno. En los ejemplos se tomará NTSC aunque los algoritmos de compresión son iguales.



* Se separan los canales Y, I y Q
* Para I y Q se toman matrices de 4 pixels (2x2) y se promedian dejando finalmente para cada uno matrices de 320x240. En esta reducción se pierde información, pero no es tan importante ya que se trata de la crominancia, en la que el ojo humano tiene menor respuesta que en la luminosidad.
* Se comprime la información usando un factor de 2
* Se resta 128 a todos los elementos de las 3 matrices haciendo que su rango esté entre (-128, 127).
* Finalmente se dividen las matrices en submatrices de 8x8:
  + Y: 4800 submatrices
  + I: 1200 submatrices
  + Q: 1200 submatrices

Paso 2:

* Se aplica DCT (Discrete Cosine Transformation – transformación discreta del coseno) a los 7200 bloques de manera separada. La salida de cada una es una matriz 8x8 de coeficientes DCT.
* El elemento (0,0) es el valor promedio del bloque. El resto dirá cuanto poder de espectro está presente en cada frecuencia espacial. DCT es similar a la transformada del espacio de Fourier de dos dimensiones

