

Tecnologías multimedia en Web

Máster en Ingeniería de Servicios y Aplicaciones Web

Juan Domingo Esteve (Juan.Domingo@uv.es)
Dpto. de Informática
Universidad de Valencia

Octubre de 2014

Tecnologías Multimedia en Web

Profesor: Juan Domingo (Juan.Domingo@uv.es)

Carga lectiva presencial: 16 h de teoría
6.5 h de prácticas
2.5 h de exposiciones

Horario: Martes, de 15:00 a 17:30, los días:
30 de Septiembre
14,21 y 28 de Octubre
5,11,18 y 25 de Noviembre
2 y 9 de Diciembre

Método de evaluación: Evaluación de las prácticas
(aprox. 20 %)
Exposición individual de un artículo o tema
(aprox. 40 %)
Examen escrito (2h, 13 de Enero, 15:00 h)
(aprox. 40 %)

Tutorías: Martes, de 10:00 a 13:30
(se pueden concertar por mail tutorías en otro momento, mañana o tarde).

Tecnologías Multimedia en Web

Bibliografía

- Multimedia signals and systems. S. Stanković, I. Orović, E. Sejdić, Springer, 2012. ([accesible UV](#)). Muy interesante.
- Multimedia analysis, processing and communications. Weisi Lin et al., Springer, 2011. ([accesible UV](#)). Util en ciertos capítulos.
- HTML5 multimedia development cookbook: recipes for practical, real-world HTML5 multimedia-driven development. D. Cruse, L. Jordan, Packt Pub., 2011. ([accesible UV](#)). No muy interesante.
- HTML5 Guidelines for web developers. K. Förster, B. Öggl, Addison-Wesley. Dos ejemplares pedidos. Muy interesante.

Introducción

Índice

- Definición de multimedia
- Elementos del multimedia I: texto e hipertexto
- Elementos del multimedia II: audio
- Elementos del multimedia III: imágenes
- Elementos del multimedia IV: video

Multimedia es la integración en un sólo documento digital de múltiples medios de transmisión de la información. El multimedia busca, entre otras cosas:

- Aumentar la cantidad y calidad de la información transmitida al receptor
- Ejemplificar lo que el texto simple establece
- Lograr una interacción dinámica y fluida con el receptor de la información

Los ejemplos más antiguos de multimedia son las ilustraciones añadidas a libros, especialmente códices:

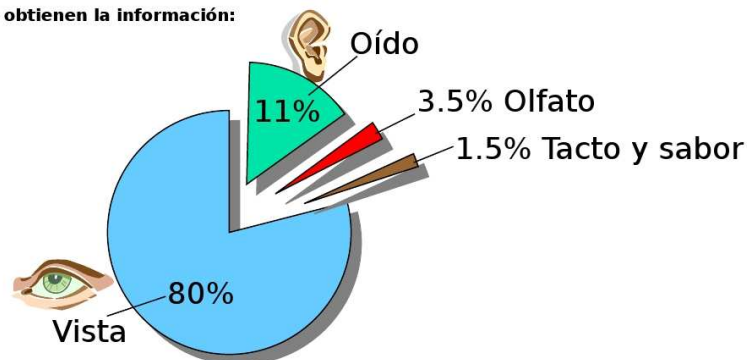


Introducción

Hay algunas ventajas obvias en el uso de multimedia para transmitir información:

- La información resulta más atractiva
- El uso de otros sentidos ayuda a la asimilación y la retentiva.

Las personas obtienen la información:



- La interactividad también favorece el aprendizaje

Campos de aplicación del multimedia:

- Producción y edición digital de video
- Periódicos electrónicos
- TV interactiva
- Video conferencia
- Móviles 3G
- Juegos interactivos
- Cursos interactivos
- Aplicaciones futuras....?

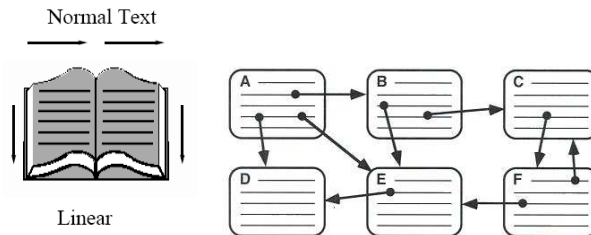
Introducción

Nos centraremos en la parte tecnológica del multimedia, en concreto, en las tecnologías informáticas usadas, que incluyen:

- Tecnologías hardware para la...
 - ... captura (micrófonos, cámaras digitales de fotografía y video...)
 - ... digitalización (convertidores A/D)
 - ... almacenamiento (discos y medios secundarios)
 - ... compresión (hardware específico)
 - ... transmisión (redes de cable o inalámbricas)
 - ... visualización y audición (pantallas, proyectores, amplificadores, altavoces...)
- Algoritmos y programas para la...
 - ... generación (modelos 3D,..)
 - ... representación (VRML,...)
 - ... procesamiento (proceso digital de señal e imagen)
 - ... compresión (codificación, transformadas integrales...)
 - ... transmisión (protocolos de Internet)
 - ... almacenamiento (bases de datos)
 - ... presentación (HTML y estándares relacionados)

El multimedia actual (el usado en documentos digitales) se integra normalmente en una estructura de **hipertexto**, casi nunca de texto simple. La diferencia es:

- Texto: sucesión de palabras de supuesta lectura lineal (secuencial)
- Hipertexto: bloques de texto simple unidos por enlaces, de lectura no necesariamente secuencial



El hipertexto en documentos digitales, aunque precedido por [Gopher](#), es creación de Tim Berners-Lee para la [primera versión](#) de HTML (HyperText Markup Language, 1990). La estructura no lineal es apropiada para documentos técnicos o científicos, o aquellos que deseen proveer exclusivamente la información deseada por el usuario de modo rápido. ATENCION: un uso incorrecto del hipertexto puede desviar la atención y distraer al usuario del contenido importante (si lo hay).

Texto en multimedia

Aunque hay varias formas de construir documentos en hipertexto, la más común es usar **lenguajes de marcado**, y en concreto, marcado de presentación, en el que se encuentra [HTML](#). Todo documento en HTML sigue la estructura de un árbol con una única raíz de la que cuelgan nodos que pueden ser otras estructuras, contenido final (posiblemente multimedia) o una referencia a otro documento HTML:

```
<!DOCTYPE HTML>
<html>
<head>
<meta charset="UTF-8">
<title>A silly example</title>
</head>
<body>
This is a text.<br/>
Now, a video example:<br/>
<video width="320" height="240" controls>
  <source src="movie.mp4" type="video/mp4">
  <source src="movie.ogg" type="video/ogg">
  Your browser does not support the video tag.
</video>
</body>
</html>
```

La única consideración relevante para el uso del texto es tener en cuenta la codificación usada:

- **ASCII**: American Standard Code for Interchange of Information. Código que asigna un número de 7 bits a cada carácter, incluyendo algunos caracteres no imprimibles. No era suficiente para acomodar los caracteres de múltiples lenguajes.
- **ISO-8859-(1 a 15)**: Norma ISO para codificación de diversos alfabetos asignando un número de 8 bits a cada carácter. Los números 0 a 127 son iguales a los correspondientes ASCII, para mantener la compatibilidad.
- **UNICODE**: Provee un número único a cada carácter sin importar la plataforma, el programa o el idioma usado. Define tres codificaciones posibles: UTF-8, UTF-16 y UTF-32. En UTF-8 y UTF-16 el número de bytes usado para codificar cada carácter está entre uno y cuatro, dependiendo del carácter. En UTF-32 es siempre cuatro.

Texto en multimedia

Rango UNICODE	Valor escalar	UTF-8
000000-00007F ⁽¹⁾	00000000 0xxxxxxx	0xxxxxxx
000080-0007FF ⁽²⁾	00000yyy yyxxxxxx	110yyyyy 10xxxxxx
000800-00FFFF ⁽³⁾	zzzzyyyy yyxxxxxx	1110zzzz 10yyyyyy 10xxxxxx
010000-10FFFF ⁽⁴⁾	000uuuuu zzzzyyyy yyxxxxxx	11110uuu 10uuzzzz 10yyyyyy 10xxxxxx
		UTF-16
		00000000 0xxxxxxx
		00000yyy yyxxxxxx
		zzzzyyyy yyxxxxxx
		110110ww wwzzzzyy 110111yy yyxxxxxx
		(www = uuuuu - 1)
		UTF-32
000000-00D7FF y 00E000-10FFFF ⁽⁵⁾	xxxxxxxx xxxxxxxx xxxxxxxx	00000000 000xxxxx xxxxxxxx xxxxxxxx

(1): Rango equivalente a US-ASCII. Codificación en un único byte donde el bit más significativo es 0

(2): Codificación en dos bytes. El primer byte comienza con 110, el segundo byte comienza con 10

(3): Codificación en tres bytes. El primer byte comienza con 1110, los bytes siguientes comienzan con 10

(4): Codificación en cuatro bytes. El primer byte comienza con 11110, los bytes siguientes comienzan con 10

(5): Codificación en cuatro bytes. Los once primeros bits son 0.

El estándar HTML5 especifica que los documentos deberán ser codificados en algún ISO, o preferentemente UTF-8 ó UTF-16, pero no UTF-32. Para webs de carácter internacional es muy recomendable usar UTF-16.

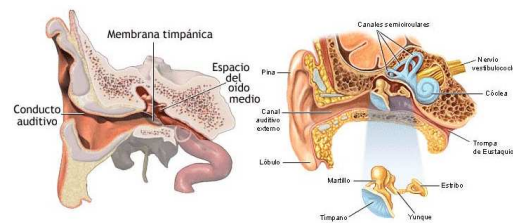
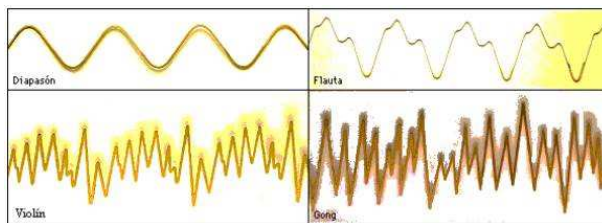
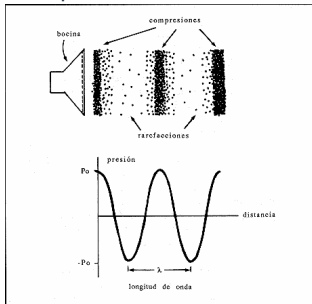
El coste de almacenamiento es menor en UTF-8 para lenguajes occidentales, pero no para lenguajes ideográficos orientales, para los que es menor el UTF-16. El coste de procesado (p. ej., ordenación alfabética) es menor en UTF-32.

Audio en multimedia

El audio se consigue digitalizando sonido. El sonido es una onda longitudinal de presión. El rango audible está entre aproximadamente 16 Hz y 20 KHz producida por la vibración de un objeto (las cuerdas vocales, una cuerda o membrana tensa) o el choque de dos objetos inmerso en un medio transmisor (aire, agua...). Sus características son:

- **Amplitud:** máximo desplazamiento de la posición de equilibrio, relacionada con la intensidad.
- **Frecuencia:** número de repeticiones (ciclos, Hz) por segundo.
- **Timbre:** diferencia en la forma de onda, característica de cada instrumento o voz.

La onda de presión llega al oído interno donde hace vibrar al tímpano, éste a la cadena de huesecillos y éstos a la cóclea, donde están los receptores nerviosos.



Tecnologías Multimedia en Web

Audio en multimedia (I)

Audio en multimedia

Las características definitorias de un sonido son:

- **Tono:** (sonido más agudo o más grave). Está relacionado con la frecuencia de repetición de la onda compleja, si ésta es periódica, y que luego identificaremos con la llamada frecuencia fundamental.
- **Intensidad:** (sonido más fuerte o más débil). Es proporcional al cuadrado de la amplitud de la onda, y por tanto proporcional a la energía transmitida por unidad de tiempo y de superficie (potencia/superficie). La sensación sonora percibida está relacionada logarítmicamente con la intensidad y por ello se suele expresar en dB:

$$I_{dB} = 20 \log_{10} \frac{P}{P_{ref}}$$

donde P es la potencia sonora recibida y P_{ref} la potencia mínima perceptible por un oído humano medio sano, la cual vale $10^{-12} W/m^2$ (ó $20 \mu P$).

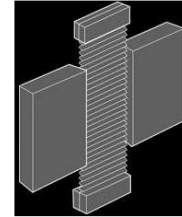
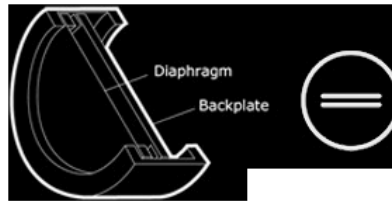
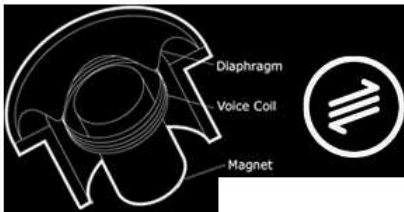
- **Timbre:** (sonido de cierto instrumento o voz frente a otro). Está relacionado con la forma de la onda, y con cómo ésta varía con el tiempo. Se entenderá mejor cuando se explique la descomposición de Fourier.

(De todos modos, los umbrales de detección y discriminación de frecuencias son muy variables; ver [más información](#))

Tecnologías Multimedia en Web

Audio en multimedia (II)

El sonido se captura con **micrófonos** que transforman la presión sonora en señal eléctrica. Pueden ser dinámicos, de condensador o de cinta:



En grabación analógica la señal se transforma en imitación en una cinta magnética (máster) que puede pasar a disco de vinilo o ser digitalizada para pasar a CD o a audio digital.

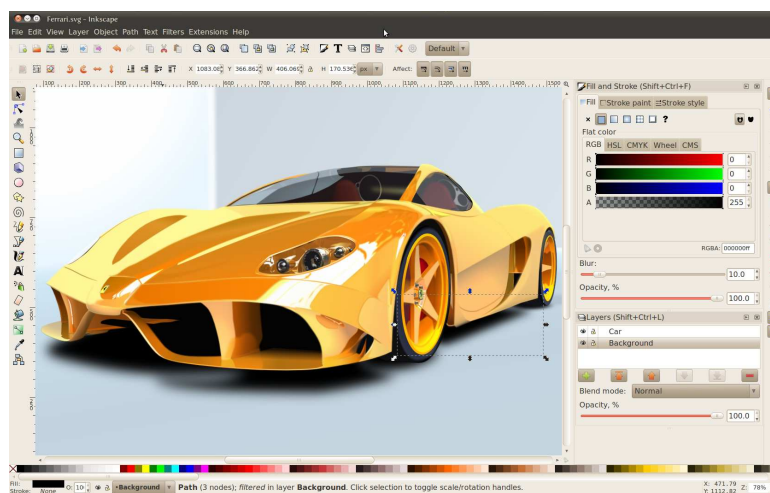
Según el número de canales (pistas) que recojamos podemos tener

- Grabación monoaural (un solo canal)
- Grabación estereofónica (dos canales, derecho+izquierdo)
- Grabación 5.1 (seis canales: frontales derecho e izquierdo, traseros izquierdo y derecho y subwoofer frontal). Está pensado para emitir sonido 3D envolvente que de impresión inmersiva (*surround*) pero es principalmente usado para audio digital.

Gráficos en multimedia

Se trata de almacenar "dibujos" que reproduzcan la realidad, y que actúen como ilustraciones. Pueden ser gráficos vectoriales en 2D o gráficos en 3D.

Los primeros están constituídos por primitivas gráficas como "punto", "línea", "polígono", "círculo", etc. con atributos que indican su localización y otros como grosor, color, sombreado, etc. La ventaja es que se les pueden aplicar transformaciones geométricas arbitrarias (reducción, ampliación, giro, deformación..) sin pérdida alguna de información. El inconveniente es que es difícil hacerlos fotorealistas.



Los gráficos vectoriales también se pueden usar para dibujar caracteres, y de hecho las fonts son originariamente gráfico vectorial. Ello hace que se puedan usar también los caracteres como primitivas.

Formatos de gráficos vectoriales:

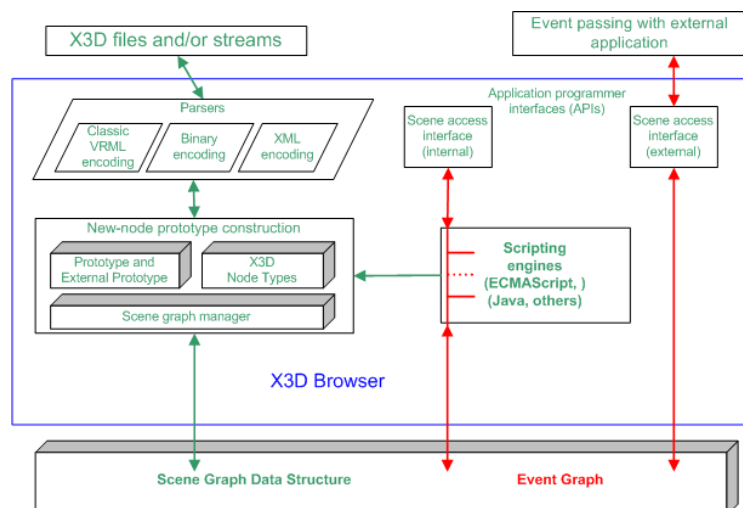
- Propios de programas comerciales
 - DWG (Autocad)
 - PostScript (Adobe), está [documentado](#).
 - PDF (Adobe Acrobat), está documentado ([ISO 32000-1:2008](#)).
 - HPGL (Hewlett Packard Graphic Language) para plotters.
- Abiertos
 - **DXF** (Drawing Exchange Format): formato estándar para el intercambio de datos entre programas de CAD
 - **SVG** (Scalable Vector Graphics): el estándar adoptado por la W3C. Es el que debería usarse universalmente de modo preferente. Los navegadores implementan trazadores de SVG.

Gráficos en multimedia

Con la misma filosofía de los gráficos vectoriales, pero usando primitivas volumétricas (ortopedro, elipsoide, cilindro, cono, figura extruída...), se pueden definir lenguajes descriptivos de gráficos en 3D. Entre ellos destacan VRML, X3D y XDOM:

VRML (Virtual Reality Modelling Language): el primero (1995) en usar esta técnica.

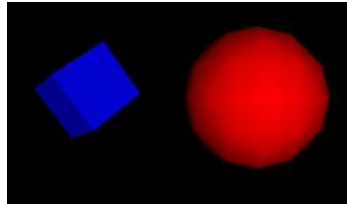
X3D: es la implementación avanzada de VRML (de hecho, es un superconjunto). Cada escena está descrita por un grafo de escena, y opcionalmente un grafo de eventos.



Un ejemplo de X3D es el siguiente:

```
#X3D V3.0 utf8
PROFILE Interchange
META filename "RedSphereBlueBox.x3dv"

Transform {
  children [
    NavigationInfo {
      headlight FALSE
      avatarSize [ 0.25 1.6 0.75 ]
      type [ "EXAMINE" ]
    }
    DirectionalLight {
    }
    Transform {
      translation 3.0 0.0 1.0
      children [
        Shape {
          geometry Sphere { radius 2.3 }
          appearance Appearance {
            material Material { diffuseColor 1.0 0.0 0.0 }
          }
        }
      ]
    }
    Transform {
      translation -2.4 0.2 1.0
      rotation 0.0 0.707 0.707 0.9
      children [
        Shape {
          geometry Box {
          }
          appearance Appearance {
            material Material { diffuseColor 0.0 0.0 1.0 }
          }
        }
      ]
    }
  ]
}
```



X3DOM: es la expresión de una descripción X3D en la forma de documento XML, lo cual ayuda a su análisis sintáctico y a su serialización para la transmisión por la web. Desde el nacimiento de HTML5 se ha incorporado en sus especificaciones.

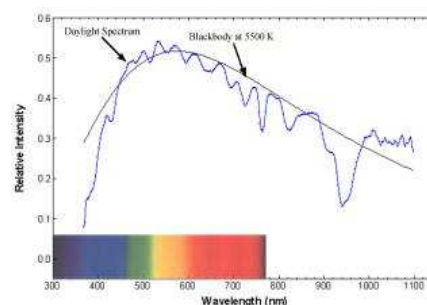
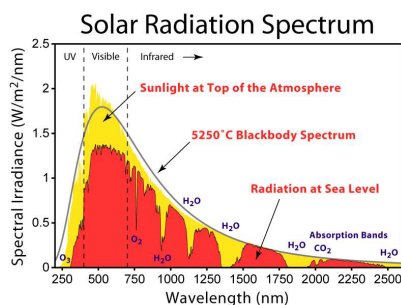
Tecnologías Multimedia en Web

Gráficos en multimedia (IV)

Imágenes en multimedia

Una imagen es una representación en dos dimensiones de objetos o escenas tridimensionales. Pueden proceder de una generación sintética, normalmente a través de descripciones geométricas como las ya comentadas, o a través de la captura de la realidad con una cámara.

A diferencia del sonido que era una onda mecánica de presión longitudinal, la **luz** que es capturada para formar una imagen es una **onda electromagnética transversal** que se propaga tanto en un medio material como en el vacío. La longitud de onda de la luz visible para humanos está entre 400 y 780 nm ($1\text{nm} = 10^{-9}\text{m}$).

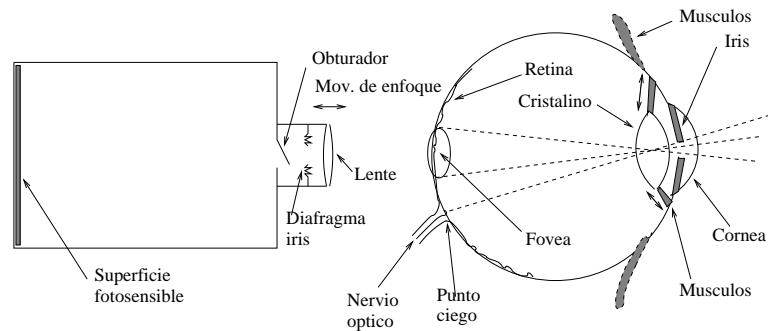


El color medido por un espectrofotómetro (**no** el percibido por humanos) viene determinado por la longitud de onda.

Tecnologías Multimedia en Web

Imágenes en multimedia (I)

El ojo humano y la cámara fotográfica.



La superficie fotosensible de la cámara es uniforme, la del ojo es distinta según la zona (fóvea y resto).

El enfoque de la cámara se hace moviendo la lente sin cambiar la distancia focal; el del ojo, deformando la lente para cambiarla.

El ojo se mueve para cambiar la dirección de su eje óptico; la cámara no (al menos, sin un trípode motorizado).

El ojo humano usa el párpado como obturador mecánico; además, controla también la intensidad de luz mediante la variación del diámetro del iris y mediante procesos

electroquímicos en la retina.

Tecnologías Multimedia en Web

Imágenes en multimedia (II)

Imágenes en multimedia

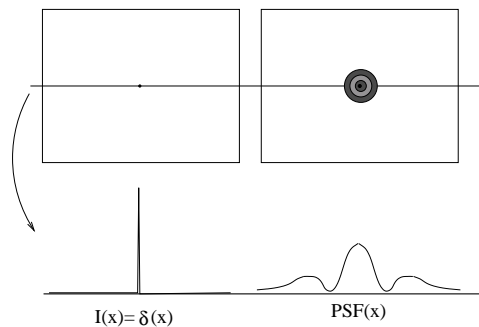
En la retina existen dos tipos de células receptoras de la luz: bastones (sensibles a la intensidad) y conos (sensibles sobre todo a ciertas frecuencias, con lo que sirven para captar el color).

Pérdidas de información en la captura, en el sistema visual humano:

- Densidad finita de células, que limita los detalles. La agudeza visual se define como $R = \frac{1}{\alpha_m}$.
- Percepción logarítmica de la intensidad, que amplía el rango pero disminuye el contraste en zonas claras.
- Tiempo finito de proceso, que limita la percepción del movimiento. No podemos distinguir imágenes individuales por encima de las 50 ó 60 por segundo.
- Proyección: todos los puntos de una misma recta se proyectan en el mismo punto del plano de proyección. No obstante, esto se suple con dos ojos.

Pérdidas de información en la captura, en sistemas artificiales:

- Proyección: como en el caso de sistemas naturales.
- Dispersión: los sistemas de captura de la imagen no generan una respuesta idéntica a la señal que los estimula, sino que la transforman. Dicha transformación se puede caracterizar usando la función de dispersión del punto (PSF), definida como la función que da el valor de la magnitud medida para una señal que consista sólo en un punto físico de extensión despreciable que emite una unidad de intensidad de la magnitud que se mida.

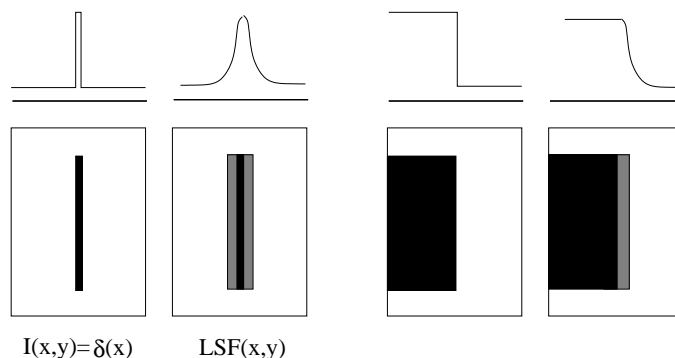


Imágenes en multimedia

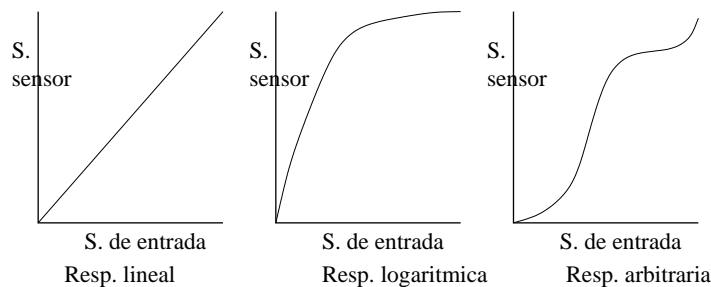
Para muchos sistemas de imagen se puede asumir que la PSF es una Gaussiana, es decir, si la fuente es $I(x) = \delta(x - x_0)$,

$$PSF(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - x_0)^2}{\sigma^2}\right)$$

Análogamente, se puede estudiar cómo se modifica la señal dada por una línea delgada o por un borde (discontinuidad entre dos regiones de diferente valor de la señal). La función de dispersión de la línea (LSF) es $LSF(y) = \int PSF(x, y) dx$



Por otra parte, la intensidad de la señal también sufre modificaciones: la señal que el sensor genera como respuesta no es estrictamente proporcional al valor de la señal de entrada. Esta modificación viene caracterizada por la curva de respuesta del sensor. Puede ser lineal (lo ideal), logarítmica (es bastante común, y aunque pierde información sobre todo en los niveles más altos permite un rango de respuesta mucho mayor) o de cualquier otra forma arbitraria (requiere conocerla mediante un calibrado del sensor).



Esto lleva a la noción de **contraste** de una imagen, que es el cociente entre el rango de valores realmente ocupado por la señal de esa imagen y el potencialmente ocupable.

Imágenes en multimedia

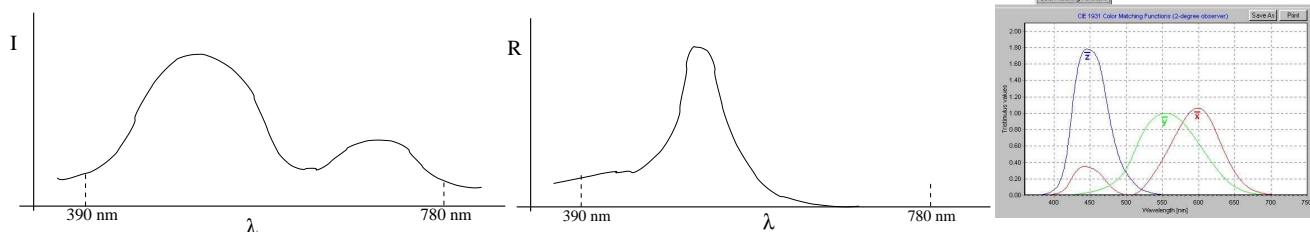
Representación del color

Se puede asignar un "color" a una luz monocromática, según su longitud de onda:

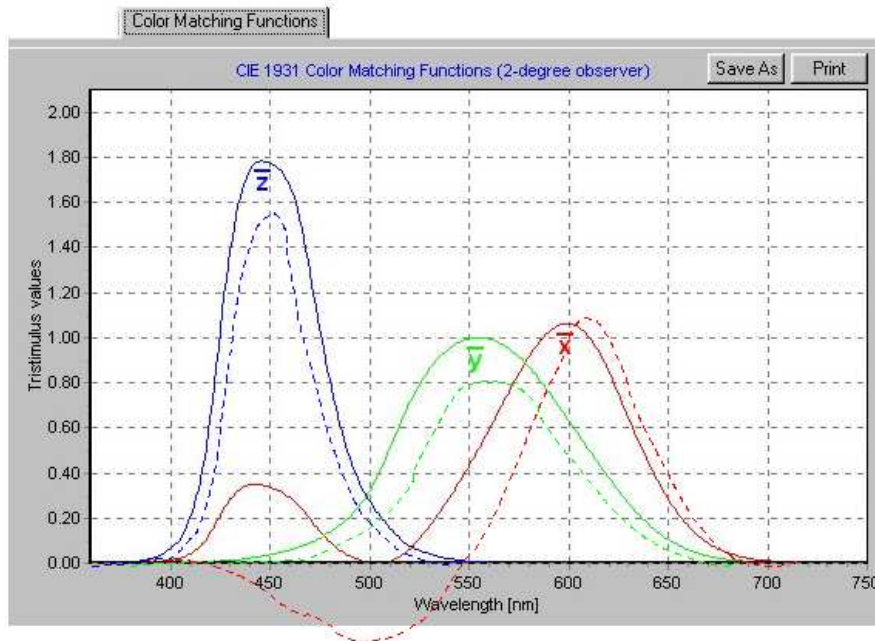
Color	λ (nm)
violeta	390—455
azul	455—492
verde	492—577
amarillo	577—597
naranja	597—622
rojo	622—780

Sin embargo, la percepción del color depende:

- De la intensidad de luz incidente sobre el objeto para cada longitud de onda, $I(\lambda)$
- De la naturaleza del objeto (proporción de intensidad reflejada para cada longitud de onda, $R(\lambda)$)
- Del receptor (qué respuesta genera ante cada longitud de onda, $S(\lambda)$).



Resultados para la respuesta del SVH
(auténticos en línea de puntos. Normalizados para evitar valores negativos
en línea continua)



Imágenes en multimedia

La respuesta α_i que se obtiene en cada sensor para el color C es:

$$\alpha_i(C) = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) R(\lambda) S_i(\lambda) d\lambda$$

puesto que el sensor suma las intensidades de todas las λ que inciden sobre él. En el caso de que la luz incida directamente sin haberse reflejado en nada,

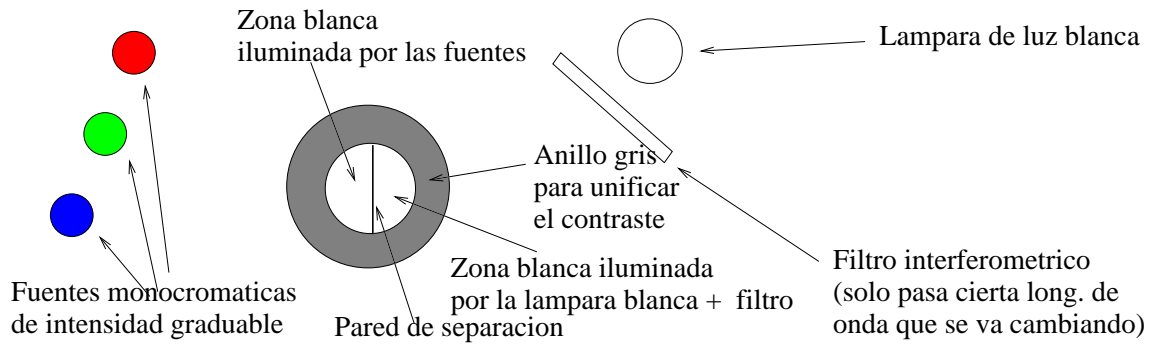
$$\alpha_i(C) = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) S_i(\lambda) d\lambda$$

Modelo triestímulo

En el SVH hay tres clases de conos, cada uno de los cuales responde con distinto valor de la señal nerviosa (electroquímica) ante cada longitud de onda.

Eso hace posible que, aunque NO se puede describir una función (la $I(\lambda)$ o la $I(\lambda)R(\lambda)$) con sólo tres valores, la suma de tres funciones con ciertos coeficientes pueda ser una aproximación razonable a la respuesta del SVH.

Medición de la respuesta del sistema visual humano:



Coordenadas CIE

Dado un iluminante, del que se conoce su función I , y tomando como funciones respuesta del sensor las S normalizadas anteriores, la función de reflectancia R (podríamos decir, el color intrínseco) de una superficie vendrá caracterizado no por su función R sino simplemente por los tres números (X, Y, Z) llamados coordenadas CIE definidas como:

Imágenes en multimedia

$$\begin{aligned} X &= k \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) R(\lambda) \overline{S}_r(\lambda) d\lambda \\ Y &= k \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) R(\lambda) \overline{S}_g(\lambda) d\lambda \\ Z &= k \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) R(\lambda) \overline{S}_b(\lambda) d\lambda \end{aligned}$$

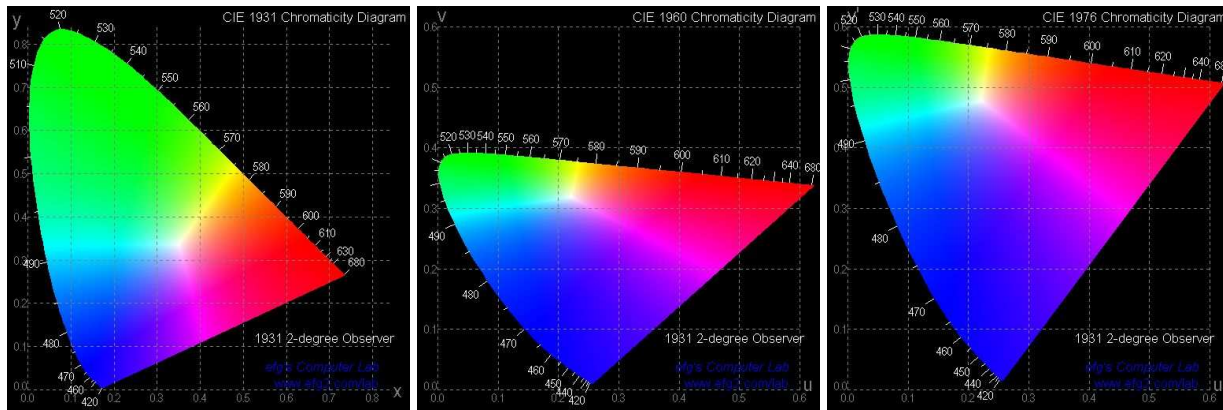
k es una constante de normalización que se escoge para que la componente Y del blanco de referencia (una iluminación fija de cierto tipo de lámpara) valga 1.

Se definen las coordenadas CIE normalizadas (x, y, z) como

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z}$$

en las que se pierde la información relativa a la intensidad global (aproximadamente igual a la suma).

Los diagramas cromáticos CIE representan los colores en el plano (x, y) . La diferencia entre unos y otros está en el iluminante escogido.



Una propiedad importante de las coordenadas CIE es que son aditivas, es decir, una iluminación formada por la unión de dos fuentes representadas por dos puntos del plano CIE se encontrará sobre la línea que las une, a distancia proporcional a la intensidad de cada componente.

Tecnologías Multimedia en Web

Imágenes en multimedia (XII)

Imágenes en multimedia

Otros sistemas de coordenadas de color

Un sistema que aproxima mejor la distancia perceptual entre colores es el (u', v') :

$$u' = \frac{4X}{X + 3Y + 15Z} \quad v' = \frac{9Y}{X + 3Y + 15Z}$$

y un sistema que sí es aproximadamente perceptual es el L^*, u^*, v^* , definido como

$$\begin{aligned} L^* &= f\left(\frac{Y}{Y_n}\right) \\ u^* &= 13L^*(u' - u'_n) \\ v^* &= 13L^*(v' - v'_n) \end{aligned}$$

siendo $f(t) = 116t^{\frac{1}{3}} - 16$ si $t > 0.008856$ y $f(t) = 903.3t$ si $t \leq 0.008856$, y (u'_n, v'_n) las coordenadas CIE del blanco de referencia.

El sistema más usado en proceso de imagen y reproducción de la señal de TV es el RGB. Se define diciendo que un iluminante I_c (la luz emitida, p. ej., por un monitor) tiene como coordenadas RGB los valores que harían que la respuesta en el SVH de una combinación de tres lámparas espectrales cuyas funciones I fueran

$$I_R = \delta(\lambda - 700.0), I_G = \delta(\lambda - 546.1), I_B = \delta(\lambda - 435.8)$$

no se diferenciase de la respuesta provocada por tal color. Es decir, se deberían igualar las cantidades c_r y c_s respectivamente para R,G y B:

$$\alpha_{Rcr}(C) = \int_{\lambda_{min}}^{\lambda_{max}} I_c(\lambda) S_R(\lambda) d\lambda$$

$$\alpha_{Gcr}(C) = \int_{\lambda_{min}}^{\lambda_{max}} I_c(\lambda) S_G(\lambda) d\lambda$$

$$\alpha_{Bcr}(C) = \int_{\lambda_{min}}^{\lambda_{max}} I_c(\lambda) S_B(\lambda) d\lambda$$

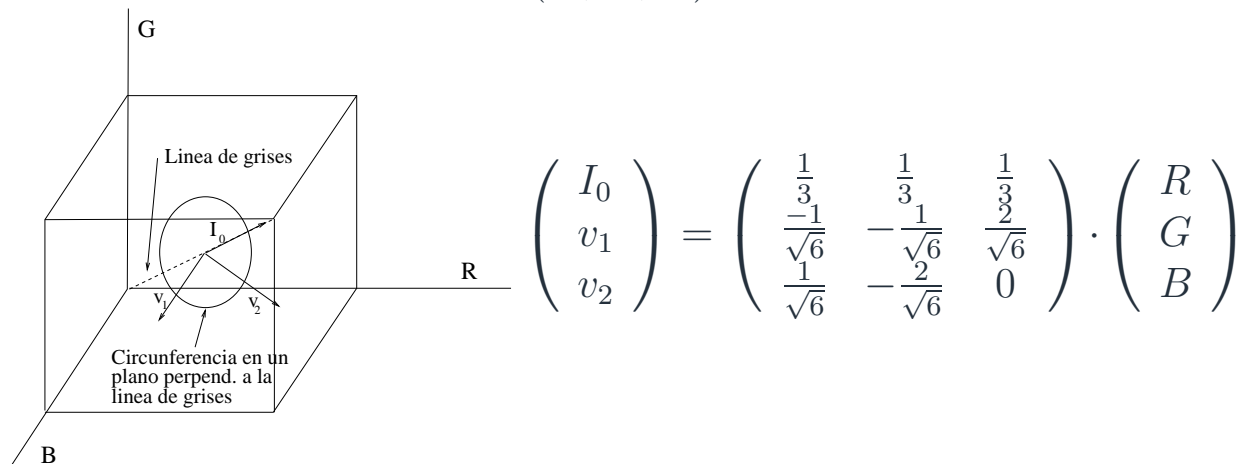
$$\alpha_{Rcs}(C) = \int_{\lambda_{min}}^{\lambda_{max}} (RI_R(\lambda) + GI_G(\lambda) + BI_B(\lambda)) S_R(\lambda) d\lambda$$

$$\alpha_{Gcs}(C) = \int_{\lambda_{min}}^{\lambda_{max}} (RI_R(\lambda) + GI_G(\lambda) + BI_B(\lambda)) S_G(\lambda) d\lambda$$

$$\alpha_{Bcs}(C) = \int_{\lambda_{min}}^{\lambda_{max}} (RI_R(\lambda) + GI_G(\lambda) + BI_B(\lambda)) S_B(\lambda) d\lambda$$

Imágenes en multimedia

Todos los colores percibibles por el SVH se puede así representar como puntos en un cubo con aristas opuestas en $(0, 0, 0)$ y $(1, 1, 1)$ del espacio (R, G, B) . A partir de las RGB se definen las (I_0, v_1, v_2) que son rotación de las anteriores:



Un sistema también ampliamente usado, sobre todo en proceso digital de imagen, es el (I, H, S) , que captura las nociones intuitivas de intensidad (la brillantez del color), saturación (cuánto se parece o diferencia un color de su correspondiente color espectral puro más cercano) y tinte (hue, que expresa el color del que se trata: rojo, naranja, amarillo, etc.). Se definen como:

$$I = I_0$$

$$S = \sqrt{v_1^2 + v_2^2}$$

$$H = \frac{1}{2\pi} \arctan\left(\frac{v_2}{v_1}\right)$$

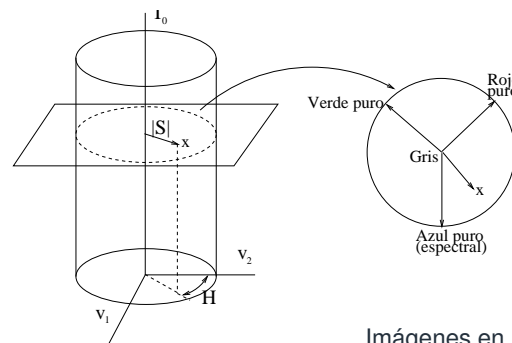
y su transformación inversa, que es

$$I_0 = I$$

$$v_1 = S \cos(2\pi H)$$

$$v_2 = S \sin(2\pi H)$$

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & -0.204124 & 0.612372 \\ 1 & -0.204124 & -0.612372 \\ 1 & 0.408248 & 0 \end{pmatrix} \begin{pmatrix} I_0 \\ v_1 \\ v_2 \end{pmatrix}$$

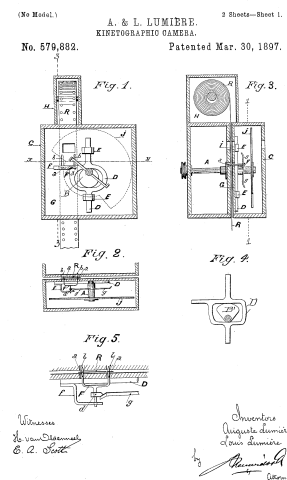


Video en multimedia

El vídeo está constituido por la sucesión rápida de imágenes que crean en el espectador una ilusión de movimiento. Se pensó que esto estaba basado en un supuesto fenómeno fisiológico llamado persistencia de la visión (POV), según el cual una imagen permanece en la retina un breve tiempo. En realidad, la ilusión de movimiento es fundamentalmente creada por el espectador, parece ser que en el área V5 del córtex visual, a través de los llamados **fenómenos phi (ϕ) y beta (β)**.

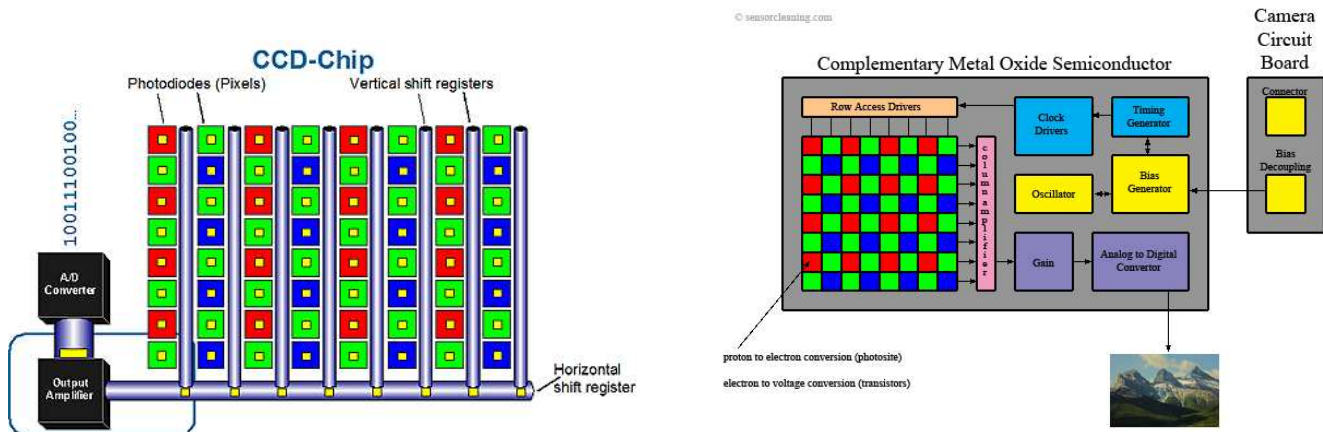
Para que se la sucesión de imágenes se perciba como movimiento continuo es necesaria una frecuencia de aproximadamente 20 o más por segundo (el cine se usan 25), y que cada una de ellas permanezca un breve instante. Esto es lo que consiguieron los hermanos Lumière con un dispositivo tipo "cruz de Malta".

patente



Cámaras

Hoy casi todas las cámaras son de estado sólido, es decir, constituídas por dispositivos semiconductores que se cargan o generan una tensión proporcional a la cantidad de fotones recibidos durante un tiempo (el tiempo de integración). Los dos tipos fundamentales son los sensores CCD y los CMOS.

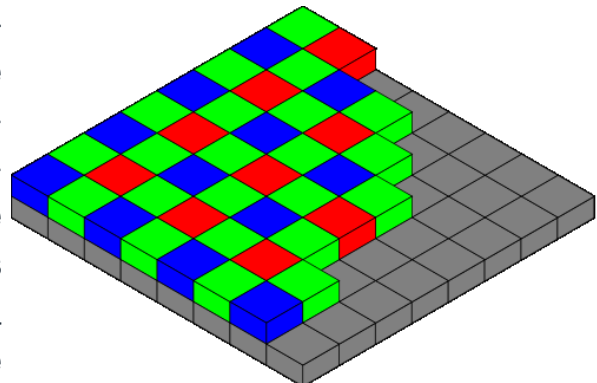


Video en multimedia

En los CCD la carga acumulada se transfiere de un elemento al siguiente y esto genera una señal que representa a una línea; la señal final es una concatenación. En los CMOS se pueden leer pixels individuales como si se tratara de una RAM. Las diferencias son:

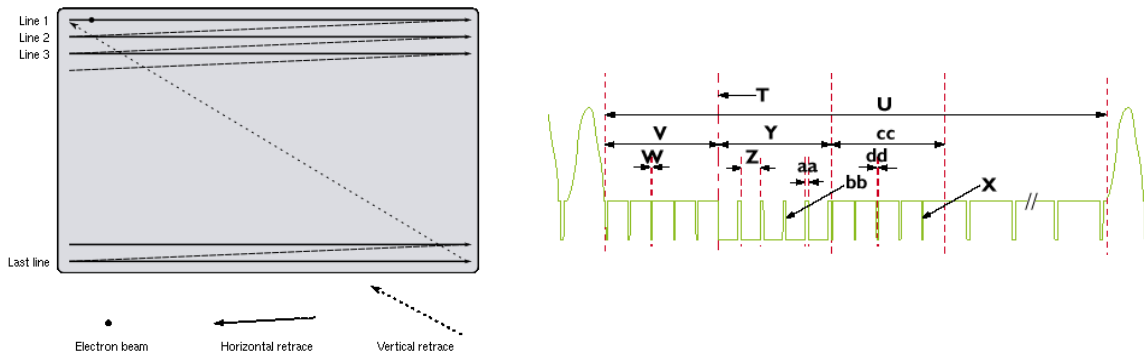
	CCD	CMOS
Sensibilidad	Muy alta	Normal
Nivel de ruido	Bajo	Normal
Potencia consumida	Muy alta	Baja
Facilidad de producción	Normal	Alta

Respecto a la captura en color, se usan elementos sensibles especialmente a las longitudes de onda del rojo, verde y azul. Como los pixels forman una retícula cuadrada y no se pueden distribuir regularmente en un mosaico tres tipos de teselas cuadradas, hay que usar más elementos de algún color, normalmente el verde, debido a la semejanza con el sistema visual humano. Esto se llama patrón o **esquema Bayer**.



Video en multimedia

El video ha sido hasta hace poco analógico. En el caso de la [televisión](#), se transmitía una señal que barría por líneas:



Cada línea indicaba el final mediante flancos (sincronismo vertical) así como cada cuadro (sincronismo horizontal). El número de líneas por cuadro varía según el estándar (525 en el americano, NTSC, 486 visibles; 625 en el Europeo, PAL, 576 visibles).

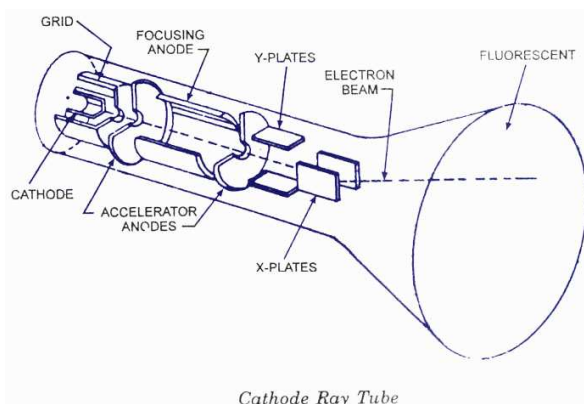
Para la televisión en color, en lugar de una señal de intensidad se transmiten tres: la Y, que coincide con la intensidad del blanco y negro (y que es $Y = 0.30R + 0.59G + 0.11B$) y dos de crominancia, que son $U = 0.493(B - Y)$ y $V = 0.877(R - Y)$.

Video en multimedia

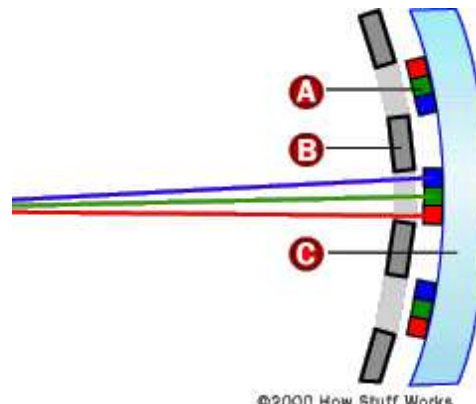
La razón de aspecto (proporción entre dimensión horizontal y vertical) es de 4:3, tanto en PAL como en NTSC

Además, ambos transmiten vídeo en formato **entrelazado**, que significa que primero se transmite un cuadro con las líneas pares y a continuación uno con las impares.

Los estándares de TV en B/N o en color estaban pensados sobre todo para ser visualizados con aparatos dotados de [tubos de rayos catódicos](#). Además, las pantallas de TV en color deben tener tres elementos fotoemisores por cada punto, y tres haces de electrones bien sincronizados.



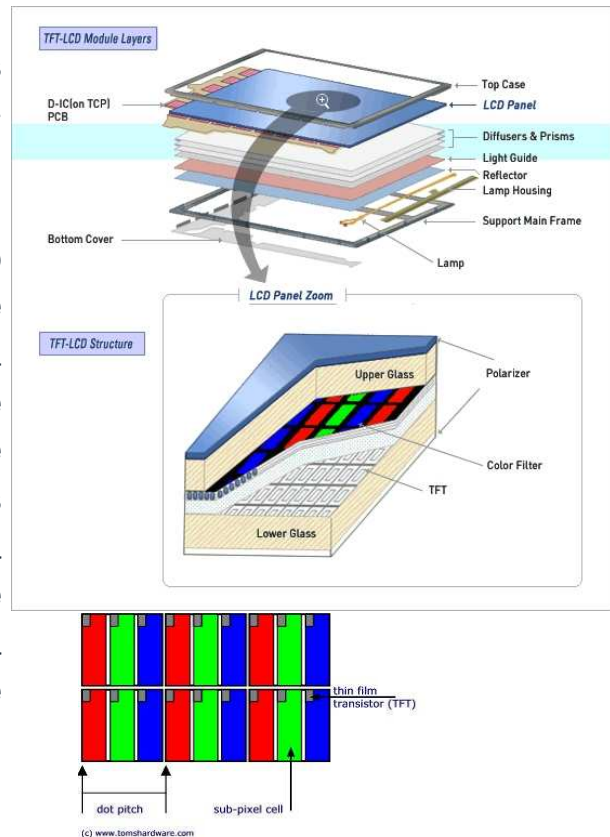
Cathode Ray Tube



©2000 How Stuff Works

Monitores

Actualmente prácticamente todos los monitores usan tecnología de estado sólido, casi exclusivamente TFT-LCD (Liquid Crystal Displays with Thin-film transistors). Se basan en un material que cambia su ángulo de polarización cuando se le aplica una tensión eléctrica, y de este modo deja pasar más o menos luz polarizada (el LCD). Para controlar cada elemento LCD se requiere construir junto a él un transistor que module la corriente (el TFT). Los materiales LCD **no** emiten luz, sólo la dejan pasar o la bloquean, por lo que es necesaria una fuente de luz posterior; ésta puede ser una lámpara de cátodo frío (CCFL, una lámpara fluorescente o una lámpara de neón) o recientemente, un conjunto de LEDs (diodos emisores de luz).



Tecnologías Multimedia en Web

video en multimedia (vi)

Video en multimedia

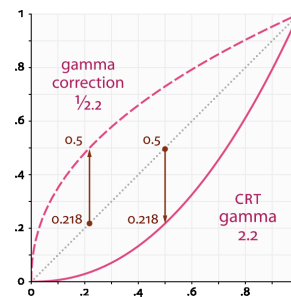
Las características principales que debemos analizar en un monitor (aparte del tamaño) son:

- Proporción H/V: tamaño horizontal del área visible entre el tamaño vertical. Suele ser 4:3 ó 16:9.
- Resolución: número de pixels en horizontal y vertical. Para un resultado óptimo se debería funcionar a la máxima resolución permisible, que correspondería a un pixel de imagen por cada elemento LCD del monitor.
- Brillantez (luminosidad): cantidad de luz emitida por unidad de área, al máximo brillo. Puede ir entre 100 y 100 cd/m^2 , y el valor óptimo está en torno a 120 cd/m^2 en ambiente normalmente iluminado.
- Contraste: el cociente entre la luminosidad máxima y la mínima que un elemento puede emitir. Desgraciadamente, la mínima no llega a ser 0 (negro absoluto), a diferencia de los monitores de tubo, en que sí es casi nula. Actualmente está entre 700:1 y llega hasta 3000:1 en monitores de grado médico.

- Tiempo de respuesta: tiempo que tarda un elemento LCD en ir desde su mínima transparencia a la máxima, y otra vez a la mínima. Es importante para la reproducción de video. Suele estar entre 20 y 35 ms.
- Espacio de color: cuántos (y qué) colores diferentes puede reproducir.
- Tasa de refresco: número de veces por segundo que puede cambiar la imagen. Está limitada por el tiempo de respuesta, y suele ir entre 60 y 75 Hz.
- Consumo energético: energía consumida por unidad de tiempo, en Watios.

En general, la retroiluminación LED mejora el contraste, amplía el espacio de color y disminuye el consumo.

Respecto a la linealidad del monitor (intensidad luminosa más o menos lineal con el valor de la señal que se le envía), la curva de respuesta suele ser de la forma $I = A \cdot s^\gamma$ donde s es el valor de la señal, normalizado en $[0..1]$, y γ es una constante, que para los monitores CRT suele estar entre 2.35 y 2.55 y para los LCD sobre 2.2.



Digitalización de señales de audio y video

Una **señal** es una magnitud física variable en el tiempo y/o en el espacio. En el caso que nos ocupa, las variables físicas de interés son la presión del aire para el sonido, y la luminancia y cada componente del color para la imagen. Como vimos, en ambos casos éstas se convertían a señal eléctrica mediante micrófonos y cámaras de vídeo.

Una señal **analógica** es aquella que toma valores en todos los instantes del tiempo, y que puede tomar cualquier valor comprendido en un intervalo conocido (p. ej., 0 a 5 Voltios).

Una señal **digital** es aquella que toma valores sólo en ciertos instantes del tiempo, llamados instantes de muestreo, normalmente separados entre sí por un tiempo constante (el periodo de muestreo, T) y además el valor es uno de entre un conjunto discreto conocido (p. ej. los enteros entre 0 y 255).

Digitalización es el proceso por el cual una señal analógica es convertida a una señal digital perdiendo información. Para ello son necesarios dos procesos: el **muestreo** y la **cuantización**.