

Sistemas audiovisuales

Apuntes de clase

Javier Rodrigo López

27 de septiembre de 2022

Índice general

1. Dispositivos de captación y reproducción de sonido e imagen	5
1.1. Micrófonos	5
1.2. Altavoces	5
1.3. Cámaras y sensores	5
1.3.1. Sistema Visual Humano	5
1.3.2. Memoria visual o integración temporal	6
1.3.3. Componentes de luminancia y crominancia	6
1.3.4. Calidad y ancho de banda de la señal de vídeo	6
1.3.5. Cámaras de vídeo	6
2. Señales y formatos de audio y vídeo	9
2.1. Digitalización de las señales de audio y vídeo	9
2.1.1. Teorema de Nyquist	9
2.2. Codificación de la señal de audio	10
2.3. Codificación de la señal de vídeo	10
2.4. Soportes y formatos de almacenamiento de audio y vídeo	10
2.5. Transmisión de señales de audio y vídeo	10
3. Introducción a los sistemas de transmisión de vídeo y audio	11
3.1. Parámetros generales de un sistema de transmisión	11
3.2. Sistemas de transmisión por cable	11
3.3. Sistemas de transmisión por fibra óptica	11
3.4. Sistemas de transmisión y difusión terrestre	11
3.5. Sistemas de transmisión por satélite	11

Presentación

Elena Blanco, José Luis Rodríguez Vázquez.

Hacer uso de las tutorías.

Capítulo 1

Dispositivos de captación y reproducción de sonido e imagen

1.1 Micrófonos

1.2 Altavoces

1.3 Cámaras y sensores

Los sistemas de vídeo van a convertir un escenario tridimensional con colores complejos en una señal unidimensional lo suficientemente parecida a la escena real como para ser entendida por el sistema visual humano.

Las frecuencias electromagnéticas que comprenden el espectro visible humano van de los 400 nm a los 750 nm. Por debajo se encuentra la radiación ultravioleta y por encima los infrarrojos.

El ojo tiene células sensibles a la intensidad luminosa (todos los colores) llamadas **bastones**, y otras células sensibles a los colores llamadas **conos**.

La **agudeza visual** es la capacidad que tiene una persona de observar dos objetos muy cercanos entre sí diferenciándolos como objetos independientes. Si adecuamos los píxeles de una pantalla, es posible dar una sensación de continuidad mediante el acercamiento de los píxeles entre sí, teniendo en cuenta la distancia a la que se encuentra el espectador.

Otra característica relacionada con la continuidad es la **memoria retentiva**. A partir de 12 o 14 imágenes por segundo, empezamos a percibir esa serie de imágenes como un vídeo continuo.

El proceso de escaneo del sensor plano (bidimensional) es lo que nos permite transformar la información en una señal unidimensional.

$$d = 42.5 \cdot \frac{D \text{ (pulgadas)}}{n^{\circ} \text{ de líneas}} \text{ m}$$

Aunque estuviera desarrollado el sistema de televisión en alta definición, los canales no habrían permitido su transmisión por falta de capacidad.

1.3.1. Sistema Visual Humano

Los conos son estimulados por diferentes longitudes de onda, habiendo tres tipos y diferenciando respectivamente el rojo, el verde y el azul. Además, no podemos distinguir si la luz proviene de una fuente única o de la superposición de varias fuentes.

Primera Ley de Grassman

Cualquier color que podemos percibir puede ser obtenido mediante la mezcla de luz roja, verde y azul (colores primarios), en una cierta proporción de intensidad.

1.3.2. Memoria visual o integración temporal

En promedio, el ojo es incapaz de distinguir varios sucesos luminosos que ocurran con un intervalo menor de 50 ms. Al comienzo de la tecnología de la televisión, algunos convenios adquirieron una frecuencia de 25 fotogramas por segundo, otros de 30. Esto sobrepasa el mínimo necesario para la sensación de continuidad, pero ayudaba a la reducción del *flickering*.

Sin embargo, la incapacidad de los canales de la época para un ancho de banda tan grande llevó a la elaboración de la tecnología de vídeo entrelazado. El vídeo era transmitido por campos (medio fotograma, conformado por las líneas pares o impares, de forma alterna) en lugar de por cuadros (fotograma entero, un solo campo). El método de barrido de cuadros se denomina **barrido progresivo**, mientras que el método de barrido de campos se denomina **barrido entrelazado**.

Estos dos métodos de barrido son importantes a la hora de diseñar las señales de sincronismo.

1.3.3. Componentes de luminancia y crominancia

Tanto en captación como en reproducción, usaremos un espacio de color RGB. Sin embargo, las señales de producción y transmisión utilizan las componentes de luminancia (brillo, corresponde con la intensidad en la televisión en blanco y negro) y crominancia. La relación que tienen con el espacio RGB es la siguiente:

$$Y = 0.2627R + 0.6780G + 0.0593B$$

$$C'_B = \frac{B' - Y'}{1.8814}$$

$$C'_R = \frac{R' - Y'}{1.4746}$$

Las componentes de crominancia valen 0 cuando la transmisión es monocromática. Por el funcionamiento del sistema visual humano, la luminancia Y nos aporta más información en la imagen que la crominancia. Por ello se estableció el doble de ancho de banda para la luminancia que para cada componente de crominancia.

1.3.4. Calidad y ancho de banda de la señal de vídeo

El ancho de banda de las señales R, G y B se estableció en 6 MHz teniendo en cuenta el número de líneas, la relación de aspecto y la frecuencia de cuadro. En componentes, la luminancia se estableció en 6 MHz y los de crominancia en 3 MHz cada uno para SD, y 36 MHz y 18 MHz respectivamente para HD.

1.3.5. Cámaras de vídeo

Las cámaras de uso profesional suelen incorporar un prisma dicróico que separa la luz blanca en las longitudes de onda correspondientes al verde, al rojo y al azul. Habrá un CCD por cada componente, obteniendo así una mayor resolución efectiva y, por tanto, una mayor calidad.

Las cámaras que no poseen esta estructura interna tienen otras formas de captar las tres señales sobre el sensor. Una muy común es el filtro Bayer, el cual está conformado por un mosaico de filtros verdes rojos y azules, de forma que cada píxel del sensor es estimulado por una sola longitud de onda. El resto de la información es obtenida mediante interpolación.

También es común encontrar filtros anti-*aliasing* colocados ante el filtro Bayer para evitar el solapamiento espectral espacial.

Los transductores de luz a señal eléctrica (los sensores) son básicamente arrays de fotodiodos. Cada fotodiodo genera una carga proporcional a la luz incidente en ese píxel. Esta información analógica es escaneada mediante barridos del sensor y posteriormente digitalizada.

Los sensores pueden ser principalmente de dos tecnologías: CCD y CMOS. Ambos son distribuidos como chips integrados que tendrán como características a tener en cuenta la sensibilidad, la resolución y la relación señal/ruido.

Ajustes de la cámara

- **Balance de blancos.** Aseguramos que el nivel máximo de señal coincide en los tres canales.
- **Balance de negros.** Aseguramos que el nivel mínimo coincide en los tres canales.
- **Ganancia.**

Sensación de parpadeo producida al ver vídeos con un bajo número de fps.

- **Nivel de negro.**

Estos ajustes pueden ser ajustados ayudándose de un vectorscopio.

Capítulo 2

Señales y formatos de audio y vídeo

2.1 Digitalización de las señales de audio y vídeo

La digitalización consiste en tres pasos fundamentales: **muestreo**, **cuantificación** y **codificación**.

El muestreo consiste en tomar valores cada ciertos intervalos de tiempo, de modo que hemos discretizado la dimensión temporal. Sin embargo, los valores de cada muestra aún son analógicos. La cuantificación consiste en asignar a cada muestra un número de un rango establecido. La codificación se refiere a la binarización de dicho número.

La cuantificación dependerá de los valores binarios que podemos almacenar. En otras palabras, del número de bits que tenga nuestro codificador, puesto que con 3 bits solo podremos cuantificar las muestras en 8 valores diferentes ($2^{3 \text{ bits}} = 8$ valores posibles). Esto significa que estamos introduciendo un error de codificación.

- **Codificación de fuente.** Compresión de la señal.
- **Codificación de canal.** Protección de la señal ante distorsiones y otros errores. Consiste en añadir bits de redundancia y otros mecanismos similares.
- **Codificación de línea.** Adapta la señal al medio, convirtiéndola en símbolos que podrán ser transmitidos por el canal (por ejemplo, una señal eléctrica).

El **audio** es una señal eléctrica que representa la variación de la presión sonora en función del tiempo.

El **vídeo** es una señal eléctrica que representa la variación de la intensidad de luz en función del espacio (imagen) y del tiempo (vídeo).

2.1.1. Teorema de Nyquist

La frecuencia de muestreo ha de ser superior al doble de la máxima frecuencia que tenga la señal a muestrear.

$$f_s \geq 2 \cdot f_m$$

Si no se cumple el teorema de Nyquist, se produce solapamiento espectral. Para solucionarlo podemos aumentar la frecuencia de muestreo o, si no es posible, introducir un filtro paso-bajo anti-*aliasing* para ignorar las frecuencias que producirán esta distorsión espectral. Los filtros reales nos imponen escoger un balance entre la calidad del filtrado (un filtro muy selectivo es más caro) y la calidad del muestreo, según nuestras necesidades. Además, el efecto apertura nos puede atenuar algunas frecuencias en función de la duración del pulso (que idealmente debería ser una delta), ya que espectralmente es una multiplicación por una función sinc.

2.2 Codificación de la señal de audio

2.3 Codificación de la señal de vídeo

2.4 Soportes y formatos de almacenamiento de audio y vídeo

2.5 Transmisión de señales de audio y vídeo

Capítulo 3

Introducción a los sistemas de transmisión de vídeo y audio

3.1 Parámetros generales de un sistema de transmisión

3.2 Sistemas de transmisión por cable

3.3 Sistemas de transmisión por fibra óptica

3.4 Sistemas de transmisión y difusión terrestre

3.5 Sistemas de transmisión por satélite
