

Tecnología de la Sociedad de la Información: Procesado de la imagen

Miguel Casas Arias



Monografía en Acceso Abierto. Libre disponibilidad en Internet, permitiendo a cualquier usuario su lectura, descarga, copia, impresión, distribución o cualquier otro uso legal de la misma, sin ninguna barrera financiera, técnica o de otro tipo.

Tecnología de la Sociedad de la Información: Procesado de la Imagen

Colección Ruta Directa a la Innovación Docente nº 105

2026 AMEC Ediciones Calle Emma Penella 6. 28055. Madrid. España.

ISBN: 979-13-88576-07-2

<https://doi.org/10.63083/lamec.2026.30.mca>

Este documento está bajo licencia Creative Commons BY-NC-ND 4.0 

Esta licencia permite a los reutilizadores copiar y distribuir el material en cualquier medio o formato, únicamente sin adaptaciones, con fines no comerciales y siempre que se cite al creador.

Índice

1 El interior de la cámara

- 1.1 Introducción al procesamiento de la imagen: elementos del cuerpo de una videocámara
- 1.2 Introducción Frame rate o cadencia:
- 1.3 Funcionamiento de un Obturador: La Velocidad de Obturación (Shutter Speed)
 - 1.3.1 Tipos de obturadores:

2 La Cadencia o Frame Rate

- 2.1 Definición de Frame o fotograma
- 2.2 Frame Rate o Cadencia
 - 2.2.1 Cadencia en el cine.
 - 2.2.2 Cadencia en Televisión
 - 2.2.3 Soluciones de incompatibilidad
- 2.3 Diferencias estéticas
- 2.4 High Frame rate (HFR)

3 Separación Tricromática

- 3.1 -Prisma dicróico
- 3.2 -Máscara Bayer
- 3.3 -Foveon:

4 Sensores

- 4.1 Definición Genérica de Sensor
- 4.2 Repaso de Plano Focal y situación del Sensor dentro de la cámara.
- 4.3 El Sensor de imagen
- 4.4 Sensores CCD y Sensores CMOS
- 4.5 Arquitectura y funcionamiento de cada tecnología: CCD y CMOS
- 4.6 Ventajas y desventajas
 - 4.6.1 Desarrollos posteriores: IT, FIT, 4T...
 - 4.6.2 Nativo, interpolación y rasterizado
- 4.7 Resolución y sensibilidad (o Rango Dinámico):
- 4.8 Nivel de Sensibilidad/Ganancia del sensor (número o valor ISO):

5 Resolución:

- 5.1 Sistemas PAL y NTSC
- 5.2 Relación de aspecto (aspect ratio)
 - 5.2.1 Relación de Aspecto del Píxel:

6 Conversión Analógico Digital A/D.

- 6.1 El Muestreo de Color y el Submuestreo de Color
- 6.2 Sobremuestreo
- 6.3 Efecto Moiré o aliasing y Filtro anti-aliasing (AA, OLBF)
- 6.4 Profundidad de color, profundidad de bits, cuantificación, cuantización o Sampling depth.
 - 6.4.1 Canal Alfa:
 - 6.4.2 El estándar de 8 bits.

6.4.3 Banding:

6.5 Gamuts:

6.6 Rango Legal y Rango Extendido,

7 Sistemas De Escaneado: “P” (progresivo) e “I” (interlaced ó entrelazado). Y sistema mixto PSF.

7.1 Ventajas y desventajas de cada barrido: resolución dinámica

7.2 Cadencias del futuro

8 Visualización de la señal

8.1 El Monitor Forma de ondas

8.2 Vectorscopio

8.3 Histograma

9 Compresión

9.1 Tipos de compresión:

9.2 Ratios de compresión:

9.3 Códecs: Media y Metadatos.

1 El interior de la cámara

1.1 Introducción al procesamiento de la imagen: elementos del cuerpo de una videocámara

El interior de una cámara de video, aunque puede variar según el tipo y modelo, se compone de varias partes esenciales que trabajan en conjunto para procesar imágenes fijas y en movimiento.

El sensor de imagen es el componente esencial que captura la luz y la convierte en señales eléctricas. Los sensores son una tecnología analógica, no digital, compuestos por materiales sensibles a las diferencias de intensidad lumínica, pero no a las diferencias cromáticas. Simplemente convierten una diferencia de intensidad lumínica (fotones) en una diferencia de corriente eléctrica (voltios). Generalmente están hechos de silicio, un material sensible a la luz. Este sensor es crucial porque transforma la luz que entra por el objetivo en señales eléctricas que luego serán procesadas.

El obturador es otro componente clave. Es un mecanismo que controla la cantidad de tiempo durante el cual el sensor está expuesto a la luz. Lo explicaremos a fondo en el siguiente apartado de los apuntes.

El procesador de imagen es el encargado de convertir las señales eléctricas del sensor en datos digitales. Realiza varias tareas esenciales como la conversión analógico-digital, el procesamiento de color, la reducción de ruido y la compresión de archivos. Este componente toma las señales eléctricas crudas provenientes del sensor y las transforma en una imagen digital que puede ser almacenada y visualizada. El procesador aplica algoritmos para mejorar la calidad de la imagen, ajustar el balance de blancos, la saturación, el contraste y otros parámetros que afectan el resultado final.

La cámara también cuenta con un visor y/o pantalla que permite al usuario ver la escena antes de capturarla, revisar las imágenes y ajustar la configuración de la cámara. El visor puede ser óptico o electrónico, y la pantalla suele ser LCD u OLED. Estos dispositivos proporcionan una interfaz visual para componer la imagen, acceder a los menús y realizar ajustes en tiempo real. La pantalla también es útil para revisar las fotos y videos capturados, permitiendo verificar si se obtuvieron los resultados deseados.

El almacenamiento es otro componente importante. Las cámaras de video suelen utilizar tarjetas de memoria, como SD, microSD o CompactFlash, donde se guardan las imágenes y los videos capturados. Algunas cámaras también pueden tener almacenamiento interno. La capacidad y velocidad de la tarjeta afectan la cantidad de material que se puede almacenar y la rapidez con que se graban los datos, especialmente en videos de alta resolución.

La batería suministra energía a la cámara para alimentar sus componentes y funciones. La duración de la batería puede variar según el modelo y el uso, especialmente si se emplean funciones que consumen mucha energía como el uso continuo de la pantalla, la grabación en alta resolución o la conectividad inalámbrica. Es esencial contar con suficiente energía para asegurar que la cámara funcione correctamente durante el tiempo necesario.

En el interior de la cámara también se encuentran los controles y botones que permiten ajustar la configuración, cambiar modos de captura y acceder a diversas funciones. Estos controles permiten al usuario modificar parámetros como la apertura, velocidad de obturación, sensibilidad ISO, balance de blancos y modos de enfoque. La disposición y cantidad de controles pueden variar

según el modelo de la cámara, ofreciendo más o menos opciones de personalización y control manual.

Las cámaras de video suelen tener micrófonos incorporados para capturar el audio durante la grabación. Algunas también pueden tener altavoces para reproducir el audio al revisar los videos. En la mayoría de modelos, es posible conectar micrófonos externos para mejorar la calidad del sonido. El sistema de audio es importante para complementar las imágenes con sonido de buena calidad, especialmente en producciones profesionales.

La estabilización de imagen es otra característica importante. Ayuda a reducir el efecto de las vibraciones y movimientos involuntarios, mejorando la calidad de la imagen. Puede ser óptica, mediante elementos móviles en la lente o el sensor, o electrónica, compensando el movimiento mediante procesamiento digital. La estabilización es especialmente útil en situaciones de poca luz (entenderemos el por qué más adelante) o al grabar video mientras se está en movimiento.

La conectividad es un aspecto cada vez más relevante en las cámaras modernas. Muchas ofrecen opciones para transferir archivos y controlar la cámara de forma remota mediante puertos físicos como USB o HDMI, y conexiones inalámbricas como Wi-Fi o Bluetooth. Esto facilita compartir las imágenes y videos, actualizar el firmware de la cámara y utilizar aplicaciones móviles para controlarla a distancia.

El software interno y firmware de la cámara gestionan todas las funciones y operaciones. Incluyen la interfaz de usuario, los menús y opciones que permiten personalizar la experiencia de uso. Los fabricantes pueden lanzar actualizaciones de firmware para mejorar el rendimiento, corregir errores o añadir nuevas funciones. El software también incluye modos de escena preestablecidos y efectos que ayudan al usuario a obtener mejores resultados en diferentes situaciones.

En cámaras de alta gama o aquellas que graban en resoluciones muy altas, el sistema de enfriamiento es importante para disipar el calor generado y evitar el sobrecalentamiento. Esto asegura que la cámara pueda funcionar de manera continua sin degradar la calidad de la imagen o sufrir daños.

1.2 Introducción Frame rate o cadencia:

Básicamente es el número de imágenes por segundo que utilizaremos en nuestra grabación de video. Aunque estudiaremos este concepto más en detalle algo más adelante, conviene entenderlo mínimamente para poder continuar explicando el siguiente apartado sobre la Velocidad de Obturación, sobretodo para identificar que hablamos de cosas diferentes y poder diferenciarlas desde un primer momento.

1.3 Funcionamiento de un Obturador: La Velocidad de Obturación (Shutter Speed)

El funcionamiento del obturador es fundamental para el control de la exposición y el efecto visual de las imágenes capturadas. Su mecanismo se puede configurar para ajustar la velocidad con que se abre y se cierra. Esto determina la cantidad de luz que recibe el sensor y afecta estéticamente la captura de movimiento y la nitidez de las imágenes. Por lo tanto, el obturador es el dispositivo de la cámara que controla el tiempo durante el cual el sensor está expuesto a la luz. Es una herramienta esencial para controlar de manera intencionada la cantidad de luz que queremos capturar. Cuanto menos tiempo permanezca expuesto el sensor a la luz, menos información

luminica recibirá, y viceversa. Y dicho de otra manera, cuanto más rápida sea la velocidad de obturación, menor cantidad de luz entrará al sensor, y al contrario.

Por ejemplo, velocidades de obturación rápidas congelan el movimiento, y son ideales para capturar sujetos en acción con movimiento sin borrosidad, mientras que velocidades de obturación lentas permiten efectos creativos como el desenfoque de movimiento, útil en situaciones de poca luz o para lograr efectos artísticos.

Las velocidades de obturación se miden en segundos o fracciones de segundo. Por ejemplo, $1/4$ significa un cuarto de segundo, mientras que $1/250$ representa un doscientos cincuentavo de segundo, que equivale a cuatro milisegundos.

El obturador de una cámara de fotos y una de video digital son tecnologías diferentes que veremos más adelante.

Llamamos entonces velocidad de obturación a la rapidez con la que se abre y cierra el obturador.

En video, la velocidad de obturación estándar es el inverso del doble de la cadencia (cuadros por segundo o frame rate) que estemos utilizando. Es decir, si grabamos video con una tasa de 25 fotogramas por segundo, la velocidad de obturación sería de $1/50$ de segundo. Esta configuración corresponde a la estética a la que estamos acostumbrados en el contenido audiovisual, aunque existen muchas excepciones con intenciones creativas específicas mediante otras configuraciones.

Una velocidad de obturación lenta puede ser útil en ciertas situaciones creativas, pero debe usarse con cuidado y ajustarse adecuadamente para lograr el efecto deseado. Si se configura incorrectamente, puede provocar imágenes borrosas y una exposición inadecuada. Por otro lado, una velocidad de obturación rápida se utiliza para congelar el movimiento y obtener imágenes nítidas en condiciones de alta velocidad. Estéticamente hablando, una velocidad de obturación muy alta produce un efecto más estroboscópico, y la imagen se ve muy definida. En cambio, con una velocidad más baja, cercana al estándar, se consigue un movimiento más "natural", con bordes difuminados que ayudan a transmitir la sensación de movimiento.

1.3.1 Tipos de obturadores:

El obturador mecánico abre y cierra físicamente una cortinilla. Es un conjunto de láminas que se desplazan por delante del sensor y con una velocidad en función del tiempo de exposición que le hayamos dado en la configuración. De esta manera, el sensor recibe la información por partes.

Los obturadores mecánicos son comunes en fotografía pero para grabar video sólo se encuentran en cámaras de cine de alta gama. Estos obturadores mecánicos pueden ofrecer un mayor control creativo en ciertas situaciones, pero también pueden ser más costosos y requieren más mantenimiento que los obturadores electrónicos.

El obturador electrónico: El obturador electrónico es un componente fundamental en las cámaras de video digitales modernas. Su función es controlar el tiempo durante el cual el sensor de imagen está expuesto a la luz, regulando así la cantidad de luz que llega a cada fotograma. A diferencia del obturador mecánico, que utiliza partes físicas para abrir y cerrar el paso de luz, el obturador electrónico realiza esta tarea mediante la activación y desactivación electrónica del sensor.

El funcionamiento del obturador electrónico consiste igualmente en controlar el inicio y el fin de la captación de luz por parte del sensor. Al comienzo de la exposición, el sensor se activa electrónicamente, comenzando a recopilar información luminica. Después del tiempo establecido, conocido como velocidad de obturación, el sensor se desactiva, deteniendo la captación de luz. Este proceso es similar a encender y apagar el paso de luz, pero a diferencia del sensor mecánico,

sin partes mecánicas móviles. Después, el sensor permanece inactivo hasta que se inicia el siguiente ciclo de exposición.

Podemos decir entonces que en el obturador electrónico, la velocidad de obturación es el tiempo durante el cual el sensor está activo.

Entre las ventajas del obturador electrónico, destaca la ausencia de partes mecánicas. Esto ofrece mayor fiabilidad, ya que no hay desgaste por movimiento de piezas, reduciendo el riesgo de fallos mecánicos. Además, su operación es silenciosa al no haber partes móviles, lo que es ideal para entornos donde el ruido debe minimizarse. También permite velocidades de obturación más rápidas, alcanzando incluso 1/32,000 de segundo o más, lo que facilita la captura de acciones rápidas con buena definición. La eliminación del mecanismo del obturador mecánico permite diseñar cámaras más compactas y ligeras.

A su vez, los obturadores electrónicos pueden ser de dos tipos según su funcionamiento:

1. Electrónico no global: (Rolling Shutter):

El obturador electrónico no global, conocido comúnmente como rolling shutter, es un método de captura de imágenes en sensores digitales donde la información se registra línea por línea, ya sea horizontal o verticalmente, dependiendo del diseño del sensor. Este proceso continúa hasta que se completa todo el recorrido del sensor, y el tiempo total que toma corresponde al tiempo de exposición establecido.

Para comprender cómo funciona el rolling shutter, podemos compararlo con el obturador mecánico de cortinilla en cámaras tradicionales. En el caso del rolling shutter electrónico, aunque no hay partes mecánicas en movimiento, el efecto es similar: los píxeles o fotodiodos del sensor comienzan a exponerse secuencialmente, línea por línea. Primero, la primera línea de píxeles se activa y comienza a recopilar luz; inmediatamente después, la siguiente línea se activa, y así sucesivamente, hasta que todas las líneas han iniciado su exposición.

Una vez que el tiempo de exposición para la primera línea ha transcurrido, esa línea deja de recopilar luz (se "cierra"), mientras que las demás líneas continúan su exposición hasta completar su respectivo tiempo. Después de que cada línea ha finalizado su exposición, se lee y almacena la información de esa línea. Es como si una cortinilla virtual pasara por el sensor, exponiendo y luego ocultando los píxeles en orden secuencial.

Debido a que no todas las partes del sensor están expuestas a la luz al mismo tiempo, sino de manera secuencial, surgen ciertas distorsiones al capturar objetos en movimiento o cuando la cámara se mueve rápidamente. Estas distorsiones ocurren porque diferentes partes de la imagen se capturan en momentos ligeramente distintos, lo que puede provocar deformaciones en la imagen.

Ejemplos prácticos:

Grabación de hélices o ruedas en movimiento: Al filmar hélices de aviones, drones o ventiladores en funcionamiento, las palas pueden aparecer curvas, retorcidas o con formas extrañas, en lugar de representarse como líneas rectas. Esto se debe a que las palas se mueven a gran velocidad, y debido a la lectura secuencial del sensor, su posición cambia entre la exposición de las diferentes líneas, creando una imagen distorsionada.

Vehículos en movimiento: Al capturar vehículos que pasan a alta velocidad, las ruedas pueden parecer ovaladas o inclinadas, y el vehículo en sí puede mostrar una deformación en su forma.

Esto ocurre porque las partes del vehículo se mueven durante el tiempo que tarda el sensor en escanear toda la imagen.

Fotografía de paneo rápido: Al hacer un paneo rápido de la cámara para seguir un objeto o capturar un panorama, los objetos verticales como edificios, postes o árboles pueden aparecer inclinados hacia un lado. La rápida rotación de la cámara hace que cada línea del sensor capture una parte de la escena en posiciones ligeramente diferentes.

Vibraciones en grabaciones de mano: Al grabar con la cámara en mano sin estabilización, especialmente al caminar o correr, la imagen puede presentar ondulaciones o temblores exagerados, dando una sensación de "gelatina". Las pequeñas vibraciones y movimientos se amplifican debido a la lectura secuencial del sensor.

Instrumentos musicales de cuerda: Al filmar cuerdas vibrando, como las de una guitarra o un violín, las cuerdas pueden aparecer como ondas o patrones inusuales en lugar de líneas rectas. Las vibraciones rápidas de las cuerdas se registran de forma desfasada en el sensor.

Para minimizar los efectos del rolling shutter, se pueden tomar varias medidas. Usar velocidades de obturación más rápidas puede reducir el tiempo de exposición, disminuyendo la distorsión, aunque esto puede requerir más luz. Estabilizar la cámara mediante trípodes, estabilizadores o cámaras con estabilización de imagen ayuda a reducir movimientos no deseados. Evitar movimientos rápidos de cámara y planificar tomas que no requieran movimientos bruscos también es efectivo. Además, algunos programas de edición de video ofrecen herramientas para corregir o minimizar las distorsiones del rolling shutter en postproducción y tarde o temprano, los desarrolladores tendrán listas herramientas de inteligencia artificial.

2. Electrónico global: /Global Shutter): El obturador electrónico global, conocido también como global shutter, es una tecnología en sensores de imagen digitales que permite capturar toda la escena de forma simultánea, exponiendo y ocultando todos los fotodiodos del sensor al mismo tiempo. A diferencia del rolling shutter, donde el sensor se lee línea por línea y puede causar distorsiones en objetos en movimiento, el global shutter elimina estos problemas al activar y desactivar todo el sensor de una vez.

Al inicio de la captura, el sensor se enciende según el tiempo de obturación configurado, y todos los píxeles comienzan a recopilar luz al mismo tiempo. Después del tiempo de exposición establecido, el sensor se apaga completamente, deteniendo la captación de luz en todos los fotodiodos simultáneamente. La información recopilada por todos los píxeles se procesa para formar la imagen final. Este proceso asegura que cada punto de la imagen se capture en el mismo instante.

El obturador electrónico global ofrece varias ventajas significativas. Primero, elimina las distorsiones por movimiento, y también posibilita una mayor tasa de fotogramas por segundo, facilitando la grabación de video en cámara lenta. El obturador electrónico global brinda control creativo de la profundidad de campo. Al utilizar altas velocidades de obturación, es posible abrir más el diafragma incluso en ambientes bien iluminados, logrando una profundidad de campo reducida y efectos estéticos deseados.

A pesar de sus ventajas, el obturador electrónico global no ha llegado ampliamente al mercado de consumo masivo. Actualmente, se encuentra principalmente en cámaras de alta gama y equipos profesionales. La complejidad y el costo de producción de sensores con global shutter han limitado su adopción en cámaras de nivel consumidor, pero se espera que esto cambie a medida que la tecnología avance.

Las aplicaciones prácticas del obturador electrónico global incluyen la cinematografía profesional, y producciones televisivas de alto presupuesto, ya que la calidad de imagen es primordial y las distorsiones por movimiento deben evitarse. También es valioso en fotografía deportiva y de acción, permitiendo capturar imágenes nítidas de sujetos en movimiento rápido sin distorsiones. Es esencial en drones y vehículos autónomos, para sistemas de visión que requieren precisión en la captura de imágenes mientras están en movimiento. Además, se utiliza en investigación y ciencia, en aplicaciones que requieren alta fidelidad en la captura de imágenes para análisis precisos, como estudios de biomecánica o pruebas balísticas.

No obstante, con la luz artificial y las altas velocidades aparecen los problemas, principalmente dos: el primero de ellos es que, al disparar en escenas con luz artificial, aparecen unas bandas en la fotografía, que delatan la frecuencia de las bombillas con las que iluminan la estancia. Y las deformaciones a altas velocidades.

2 La Cadencia o Frame Rate

2.1 Definición de Frame o fotograma

El fotograma o frame es una de las cientos o miles de imágenes que conforman una animación, una película o un video. Antiguamente, cuando la palabra fotograma surgió, esta hacía referencia a cada una de las imágenes impresas químicamente en la tira del celuloide de una película cinematográfica.

2.2 Frame Rate o Cadencia

El número de imágenes por segundo, también conocido como fotogramas por segundo (fps) o imágenes por segundo (ips), es lo que denominamos cadencia o framerate. Este parámetro es fundamental en la calidad y fluidez de la representación visual, ya que, a mayor cantidad de imágenes por segundo, mayor será la fidelidad y realismo con que se percibe el movimiento. Si el número de imágenes es insuficiente, puede producirse el efecto conocido como "parpadeo" o saltos entre fotogramas, lo cual afecta negativamente la experiencia visual del espectador.

Es importante señalar que la cadencia ha sido históricamente un compromiso entre calidad y economía. Los análisis técnicos indican que la cadencia ideal para una representación verdaderamente "real", sin sensación de parpadeo, estaría en torno a los 72 fps. Sin embargo, por razones económicas y tecnológicas, tradicionalmente no se ha ofrecido esta cantidad de fotogramas por segundo. Actualmente, ningún estándar de cine o televisión alcanza este número, aunque no se descarta que en el futuro pueda lograrse. La tecnología digital actual lo permite, y en el ámbito digital, los costos no se incrementan tanto como ocurría con los sistemas fotoquímicos.

Aunque hablamos de sistemas estandarizados, existen obras audiovisuales que utilizan tasas de fotogramas más altas. Un ejemplo claro son los videojuegos, donde es común encontrar framerates superiores a los 60 fps para lograr una experiencia más fluida y cercana a la percepción humana del movimiento. Algunas producciones cinematográficas experimentales también han explorado framerates más altos, buscando mejorar la inmersión y la calidad visual.

2.2.1 Cadencia en el cine.

En cine tradicional no digital, el compromiso viene dado, una vez más, por el coste del material. Rodar con cadencias muy altas significa aumentar proporcionalmente el gasto en película. De la misma manera, las ya de por sí voluminosas copias positivadas (cuatro o cinco rollos de gran tamaño) también aumentan, complicando su distribución en las salas.

La recomendación actual para el cine digital de la norma DCI mantiene la cadencia tradicional de 24 fps (DCI es un estándar en la industria del cine digital que establece las especificaciones técnicas para la proyección de películas digitales en salas de cine). Pero con la posibilidad de aumentarla a 48 en los formatos 2K y 4K. La cadencia de 48 fps en formatos 4K puede resultar un poco excesiva por el gran tamaño de los ficheros, pero una vez más no es descartable en un futuro a medio o largo plazo.

2.2.2 Cadencia en Televisión

El compromiso en el caso de la televisión no es tanto de calidad/coste como de calidad/flujo de datos. Dado que el "material virgen" televisivo (disco duro) no es excesivamente caro en comparación con el del cine, no supondría mucho problema de coste aumentar la cadencia. Pero sí de ancho de banda o flujo de datos: si doblamos la cadencia, también doblamos la cantidad de información.

Por razones tecnológicas, no comerciales, los sistemas de televisión optaron por cadencias similares pero un poco diferentes a las del cine. En concreto, en los sistemas PAL y SECAM se optó por los 25 fps, y en los sistemas NTSC por una cadencia un poco extraña, 29,97 fps. La razón de esta divergencia tiene que ver con la frecuencia de la red eléctrica en los distintos continentes. En Europa, la red de corriente alterna circula a 50 Hz, mientras que en Estados Unidos de América circulaba a 60 Hz. En los primitivos sistemas de televisión, esta frecuencia se usaba para sincronizar la señal, de ahí su utilidad.

Con la tecnología digital actual, la frecuencia de la corriente alterna es indiferente, pues ya no se precisa para una sincronización de la señal. Por esa razón, dado que los sistemas HD vienen a superar las viejas diferencias entre NTSC y PAL, sería conveniente que se estableciera una única cadencia para todas las producciones, bien sea 24, 25 o 30/29,97. No obstante, hay implicaciones comerciales y de compatibilidad con los sistemas antiguos que hacen que, a día de hoy, esta cadencia universal no esté definida. Por ello, la inmensa mayoría de los equipos actuales permiten, además del trabajo multiformato en SD, HD y 4K, escoger entre todas las diferentes cadencias admitidas.

2.2.3 Soluciones de incompatibilidad

Los sistemas de edición de vídeo que se utilizan hoy en día ofrecen cierta flexibilidad a la hora de definir la frecuencia de fotogramas de una secuencia o un proyecto. Las cámaras digitales pueden grabar con cualquiera de los valores de frecuencia que hemos visto (PAL, SECAM) y los sistemas de postproducción ofrecen herramientas para hacer las conversiones que puedan ser necesarias. Cuantos más fotogramas se tomen de una acción, más información se registrará sobre el movimiento que produce. Por ello la frecuencia de fotogramas por segundo de una secuencia se denomina también resolución temporal (temporal resolution).

Si grabo por ejemplo a una velocidad 50 imágenes por segundo y la reproduzco a 25 fps obtengo una cámara lenta (slow motion). Esto es diferente a grabar y reproducir ambas cosas a 50fps, que serviría para reproducir a mayor calidad los movimientos de la escena.

2.3 Diferencias estéticas

La resolución temporal en la grabación afectará a la definición de los bordes de los objetos en movimiento y al efecto de desenfoque de movimiento (motion blur). Con mucha resolución temporal supone estéticamente una diferencia notable, ya que el espectador promedio está acostumbrado a ver en el cine grabado con el desenfoque de movimiento o motion blur.

2.4 High Frame rate (HFR)

Cuando se amplía el tamaño de representación, es decir, para pantallas de mayor tamaño, es conveniente también aumentar el número de fotogramas por segundo para obtener mayor nitidez

y reproducir fielmente el movimiento. Las frecuencias de 48 fps o superiores se consideran HFR (High Frame Rate).

La frecuencia HFR produce a veces una sensación más fría (aunque teóricamente más real) que puede no agradar a todo el mundo. Pero los gustos del público también evolucionan con la tecnología y la calidad de la representación se está imponiendo con el tiempo.

La norma BT-2020 contempla, además de los 48 fps, frecuencias de 50 y 60 fps en exploración progresiva. Los proyectores de cine digital bajo la norma DCI también incluyen la posibilidad de proyectar en las salas de exhibición a 48 fps. Los nuevos desarrollos para la televisión Ultra HD consideran valores de resolución temporal más altos aún. En las recomendaciones de DVB y SMPTE está previsto alcanzar los 120 fps. Las cámaras de cine profesional e incluso alguna de gama inferior ya ofrecen también la posibilidad de aumentar la resolución temporal. Las frecuencias más habituales en difusión son: 25 fps, 29,97 fps, 24 fps, 23,98 fps y 48 fps.

3 Separación Tricromática

La separación tricromática es un proceso fundamental en la captura y reproducción del color en sistemas digitales de imagen, como cámaras de cine y televisión. Como mencionamos en la primera parte del curso, para reproducir colores utilizamos el sistema de colores primarios RGB (rojo, verde y azul). Para representar electrónicamente los colores captados del mundo real, es necesario determinar la cantidad exacta de rojo, verde y azul que tiene el objeto que queremos capturar con la cámara.

Para lograr esto, debemos descomponer o dividir la luz que entra por la lente, la cual es pancromática por naturaleza (capta todas las longitudes de onda visibles), en sus tres componentes de color primarios. Este proceso permite que cada uno de los colores primarios sea capturado y procesado de manera independiente, resultando en una reproducción precisa y fiel del color en la imagen final.

Actualmente, existen dos tecnologías principales que se utilizan en cine y televisión para este cometido: el prisma dicróico y la máscara Bayer. El prisma dicróico ha sido durante años el método más común. Por otro lado, la máscara Bayer se está imponiendo cada vez más, especialmente en las soluciones de gama alta del cine digital. Aunque ambos sistemas son considerados de calidad similar, presentan diferencias en su funcionamiento y en el procesamiento requerido.

Además de estos dos métodos, existe un tercer sistema conocido como Foveon, que aún no ha sido plenamente implementado. El sensor Foveon captura los tres colores primarios en cada píxel utilizando capas de detección apiladas que absorben diferentes longitudes de onda de luz. Esto permite una captación directa y precisa del color sin necesidad de interpolación. Aunque promete mejoras en la calidad y fidelidad del color, su adopción ha sido limitada hasta ahora debido a desafíos técnicos y costos de producción.

3.1 -Prisma dicróico

El prisma dicróico es un conjunto de filtros y espejos que se sitúa después de la lente y antes de los tres sensores captadores (las cámaras con este sistema precisan de 3 sensores de imagen).

Los filtros dicróicos tienen la particularidad de filtrar las ondas de luz por su longitud (colores) dejando pasar determinadas longitudes y reflejando otras; los espejos redireccionan los haces hacia los otros sensores.

La característica principal de este sistema es que exige un sensor diferente por cada color, pero por esa misma razón precisa sensores de menor tamaño individual que un sistema con un único sensor (con máscara Bayer como veremos a continuación).

El diseño de este prisma debe ser muy preciso, pues los tres sensores deben tener el mismo plano de enfoque con respecto a la lente.

3.2 -Máscara Bayer

El nombre se toma de su inventor, el ingeniero de Kodak Bryce E. Bayer, que inventó esta solución para poder trabajar con un único sensor. En este caso, la división cromática no se realiza por prisma, sino que se utiliza un único sensor donde se ha incrustado una máscara compuesta por “microfiltros” de los colores primarios. Esto hace que a cada píxel del sensor sólo llegue un componente primario de la luz.

El mosaico de Bayer está formado por un 50% de filtros verdes, un 25% de rojos y un 25% de azules. La razón de que se use mayor cantidad de puntos verdes es que el ojo humano es más sensible a ese color.

Aparte de los filtros en un sensor hay otros elementos, un filtro de paso bajo, un filtro de infrarrojos, microlentes y dependiendo de las cámaras unos filtros de densidad neutra.

Ventajas y desventajas

Ambos sistemas, el de máscara Bayer y el prisma dicróico, son válidos y ofrecen una calidad de imagen adecuada en la captura digital. Cada fabricante elige uno u otro basándose en motivos técnicos o comerciales. Sin embargo, es importante destacar que un sistema que utiliza un solo sensor con máscara Bayer siempre será más económico que uno que emplea tres sensores con prisma dicróico. Esto se debe a que requiere menos componentes, consume menos energía y ocupa menos espacio, lo que permite diseñar cámaras más pequeñas y ligeras.

Tradicionalmente, en televisión se han utilizado cámaras con tres sensores y prisma dicróico. Esta tecnología está muy consolidada en la industria. No obstante, uno de los problemas asociados a este sistema es el calentamiento de la cámara. El calor generado puede introducir ruido indeseado en la imagen y afectar la calidad final. Al aumentar la resolución de SD (definición estándar) a HD (alta definición), se requieren sensores más grandes, lo que incrementa aún más el calor producido. Este aumento de temperatura puede causar problemas como la descolimación o desajuste en el backfocus, debido a la dilatación térmica de las partes metálicas de la montura del objetivo.

Por estas razones, especialmente por el exceso de calor y sus efectos colaterales, la opción de utilizar un solo sensor con máscara Bayer se está extendiendo ampliamente. Este sistema reduce el calentamiento y permite construir cámaras más compactas y eficientes en términos energéticos, sin sacrificar la calidad de imagen.

3.3 -Foveon:

En este caso, la separación cromática se realiza "en vertical". No existe ningún tipo de máscara Bayer ni microfiltro de color. En el chip Foveon, cada capa es sensible a uno de los 3 colores RGB.

4 Sensores

Hoy veremos cómo el sensor de la cámara hace posible transformar la información de la onda electromagnética que entra por el objetivo (en forma de fotones), a electricidad (en forma de voltios).

4.1 Definición Genérica de Sensor

Un sensor es un dispositivo que detecta el cambio en el entorno y responde en forma de salida con la información traducida a otro sistema (convierte un fenómeno físico en un voltaje analógico medible o, a veces, una señal digital). Por ejemplo variaciones de luz, temperatura, presión, sonido, etc.

4.2 Repaso de Plano Focal y situación del Sensor dentro de la cámara.

Una vez atravesada la lente del objetivo y el dispositivo de separación tricromática, la luz llega por fin al sensor. Los sensores se encuentran en el llamado plano focal de las cámaras. El plano focal de una cámara es el punto en el que se forma la imagen capturada por la cámara. En cámaras tradicionales de película ya sea una cámara de cine o fotográfica, el plano focal es una superficie plana en la que se coloca el material fotosensible (la película). En cámaras digitales, el plano focal es un sensor de imagen que registra la luz incidente y la convierte en información electrónica, que más tarde (en otro procedimiento posterior) pasará a información digital. El sensor electrónico por tanto sustituye a la antigua película cinematográfica donde quedaba fijada la luz.

4.3 El Sensor de imagen

El sensor de imagen es el elemento de una cámara electrónica, tanto de vídeo como de fotografía, que detecta y captura la información de la luz que compone la imagen.

El sensor dispone de elementos fotosensibles (sensibles a la luz) en forma de rejilla (colocados en filas y columnas). Es una matriz de cavidades llamadas fotodiodos, que capturan la luz. Estos nos proporcionan cada una de las muestras de la realidad que pretendemos representar conjuntamente. Cada una de estas muestras es lo que posteriormente, cuando trasladamos la información recogida por la cámara y la trasladamos a una pantalla, se conoce como píxeles.

EL SENSOR ES SENSIBLE A LA INTENSIDAD DE LUZ, NO AL COLOR.

Una vez que el sensor captura la información de intensidad de luz a través de estos filtros, el procesador de la cámara realiza una interpolación para reconstruir una imagen en color completa. Este proceso implica combinar los datos de color de los píxeles adyacentes para crear una estimación del color de cada píxel, logrando así una imagen final en color. Lo veremos en clase más adelante.

El número de los fotodiodos del sensor es lo que determinará la “resolución nativa” de una imagen. Estos se convertirán en la misma cantidad de píxeles de nuestra imagen digital final siempre que no hagamos una conversión.

Todos los sensores de imagen se basan en la misma premisa: aprovechar las particularidades de materiales semiconductores (como una aleación de silicio) que pueden convertir la energía presente en la luz (fotones) en corriente eléctrica (electrones).

Una gran intensidad de la luz implica un mayor número de fotones y, por lo tanto, esto se traduce en una mayor corriente eléctrica en forma de electrones. Por el contrario, en situaciones con baja luminosidad se emiten pocos fotones y por tanto contaremos pocos electrones a la salida del sensor.

La diferencia principal entre unos sensores u otros está en la cantidad de fotodiodos que tengan y en el tamaño. Esto quiere decir que un fotodiodo podrá captar más fotones en función de su área o tamaño. El resultado es que contra más grande sea el sensor, más información y matices logrará captar.

4.4 Sensores CCD y Sensores CMOS

En su origen, las cámaras de televisión trabajaban con tubos de imagen, una tecnología actualmente en desuso. Después se desarrollaron los primeros sensores planos, con tecnología CCD (dispositivo de doble carga, en inglés). Permitían un uso más robusto de las cámaras, más eficiente que los tubos. Posteriormente, los CMOS (semiconductor complementario de metal-óxido) aparecieron en otros ámbitos de la industria, y, en los últimos años, se aplican también con éxito a soluciones profesionales de televisión y cine.

Si bien tradicionalmente se consideró el CCD de mejor calidad, hoy en día los CMOS se están imponiendo y ofrecen ventajas adicionales. Una de las ventajas principales es el menor coste de fabricación. No obstante, en un mercado en continua evolución, ambas tecnologías actualmente son plenamente válidas. Siempre ha habido, y habrá, defensores de una y otra tecnología, pero actualmente ambas se encuentran a la par en cuanto a prestaciones se refiere. Es necesario conocer, eso sí, su arquitectura, pues las dos tecnologías tienen sus ventajas y desventajas.

4.5 Arquitectura y funcionamiento de cada tecnología: CCD y CMOS

Como decíamos antes, la diferencia entre los CMOS y el CCD no son sus materiales (ambos usan el silicio como base), sino la arquitectura. Su funcionamiento determina los parámetros y el tipo de obturación que vayamos a usar. Ambas tecnologías pueden optar por máscara Bayer o por prisma dicróico para la separación de colores.

CCD: En el CCD (dispositivo de carga acoplada, en inglés charge-coupled device), el fotodiodo se usa tanto para capturar los fotones como para transferir la carga de electrones generada. Una vez expuesto a la luz, un fotodiodo descarga su carga de electrones al inmediatamente inferior. Es decir, hay una descarga por columnas. El primer fotodiodo se descarga, y, al producirse un vacío, la carga de electrones superior lo llena, produciéndose un "efecto cascada" (en lo que se conoce como "intervalo vertical").

Esta carga finalmente llega a un codificador y se transfiere a los siguientes procesadores de la señal que finalmente convertirán la señal en digital y a color.

CMOS: (Complementary Metal-Oxide-Semiconductor). Este tipo de sensor tiene un codificador en cada uno de los fotodiodos que lo componen. Es decir que es dentro del fotodiodo donde se codifican los electrones recibidos. Para ello, cada fotodiodo tiene su propia salida que los transporta directamente hasta una columna de amplificación independiente del resto de la columna.

En términos generales:

La mayoría de los sensores CCD tradicionales usan global shutter, es decir, todos los píxeles se exponen y capturan la imagen simultáneamente.

La mayoría de los sensores CMOS convencionales utilizan rolling shutter, donde la captura se hace de forma secuencial (fila a fila o columna a columna).

Pero no es una regla fija:

Existen sensores CMOS con global shutter.

Gracias a avances en el diseño de circuitos, algunos CMOS modernos incorporan memoria temporal en cada píxel, lo que les permite almacenar la carga luminosa y leerla más tarde de manera sincronizada, como un CCD. Estos sensores se usan en cámaras de alta velocidad, en aplicaciones científicas, industriales o de visión artificial, donde las distorsiones por movimiento son inaceptables.

También existen sensores CCD con mecanismos similares a un rolling shutter.

Aunque es menos frecuente, algunos CCD de bajo coste o de diseño particular pueden leer las líneas de forma secuencial, generando un efecto similar, aunque sin ser un rolling shutter en sentido estricto.

4.6 Ventajas y desventajas

Dado que el CMOS precisa espacio para otros transistores, a igual tamaño de fotodiodo, el CCD tendrá un mayor rango dinámico y menor ruido inherente.

-Consumo y tamaño: En un sensor CMOS, la luz se convierte en el propio fotodiodo, sin necesidad de otros procesadores unidos al sensor. Con esta tecnología debido a que la señal de luz no tiene que ser transportada al borde del chip para su conversión, un sensor CMOS requiere menos potencia para funcionar. Esto implica un menor tamaño, menor consumo y menor calentamiento. En este aspecto, los CCD pueden resultar más sensibles al ruido producto de ese calor.

-Producción: Los CCD son menos tolerantes con los errores de fabricación. Un único fotodiodo puede estropear el rendimiento de toda una columna. Por el contrario, los CMOS pueden ser operativos con más de un elemento defectuoso. Esto facilita su diseño y fabricación, abaratando el coste.

-Efecto shutter y artifacts: En la mayoría de los CMOS que se usan en la industria no es posible realizar toda la descarga de información de todos los fotodiodos del sensor al mismo tiempo. Su arquitectura obliga a realizarlo secuencialmente, por líneas o columnas, como si bajáramos una persiana. En caso contrario, se colapsarían los buses intermedios de almacenamiento de datos. La "captura de fotones + salida de electrones + reseteo" es lo que se conoce como "tiempo de integración". Esto supone que, si hablamos de imagen en movimiento, la captura de un CMOS nunca será perfecta e indiscutiblemente "progresiva". Hay diferencias de microsegundos entre las líneas, ocasionando los desajustes (artifacts) conocidos como "distorsión horizontal" (skew), "efecto gelatina" (jelly effect) o "burbujeo" (bubble), según hablemos, respectivamente, de lo que afecta a una imagen, al conjunto de ellas en movimiento horizontal o con una combinación de movimiento vertical y horizontal.

En el caso de los CMOS, para evitar la distorsión se intenta que la tecnología del Global Shutter se vaya implementando.

Los sensores CCD con rolling shutter a menudo se encuentran en cámaras más antiguas y algunos dispositivos de imagen científica o industrial.

-Contaminación lumínica (smear): Los CCD sin embargo están afectados por lo que se conoce como "contaminación lumínica" (smear). Esto sucede cuando un foco de luz muy intensa se dirige directamente al sensor. Las celdillas se cargan de electrones, desbordándose incluso, y afectando al resto de las celdillas de su columna. Aparece entonces en la imagen una línea horizontal sobre toda la columna. Este tipo de artifacts no aparecerán nunca en un sensor CMOS.

El antiguo sensor CCD tiene ciertos límites de tecnología y con las crecientes demandas de resolución el CCD ha llegado a los límites. Sin embargo, la tecnología CMOS, claramente tiene más espacio para el desarrollo futuro.

4.6.1 Desarrollos posteriores: IT, FIT, 4T...

En el caso de los CCD, se desarrollaron dos tecnologías, conocidas como IT (Interline Transfer) y FIT (Frame Interline Transfer). En la primera de ellas, junto a cada columna de fotodiodos se habilita otra a la que se transfiere la carga.

Esto limita (pero no evita totalmente) la contaminación.

4.6.2 Nativo, interpolación y rasterizado

Diremos que un sensor es "nativo" cuando el número de píxeles que contiene equivale (o supera) al del número de píxeles que requiere el formato final de salida de nuestro audiovisual.

En caso contrario, si el sensor (por ejemplo HD) no es nativo, pero nos ofrece una resolución de 1.080 (resolución correspondiente a una imagen HD), la conclusión es que la cámara está interpolando la resolución. Interpolación quiere decir que, mediante procesos electrónicos internos, la cámara "se inventa" un número de líneas que no son las que realmente le proporciona el sensor. Este proceso es muy común, sobre todo en cámaras domésticas pues abarata el coste, aunque en ocasiones se utiliza incluso en equipos superiores. A pesar de que la interpolación es algo negativo, los resultados pueden ser funcionales. Los mayores inconvenientes suelen venir por un aumento del ruido (sobre todo con malas condiciones lumínicas) y una falta de definición en planos con objetos que presentan líneas verticales u horizontales muy definidas (aliasing y/o moaré), y en general cualquier objeto que precise nitidez y detalle, como en planos generales de espacios abiertos.

La interpolación es una solución de alguna manera similar pero diferente del rasterizado. El rasterizado es un submuestreo que utiliza el "truco" de considerar una relación de aspecto del píxel diferente a la cuadrada 1:1, "ensanchando" la imagen; mientras la interpolación se inventa nuevas "muestras". Existen cámaras en las que a una resolución rasterizada se le añade, además, interpolación.

4.7 Resolución y sensibilidad (o Rango Dinámico):

Se llama full-frame al sensor cuyo tamaño coincide con el del negativo de 35mm.

Una vez que se conocemos el funcionamiento del sensor, lo primero que tenemos que pensar es que un este determinará dos parámetros muy importantes de una imagen digital:

1. Su resolución (o número total de píxeles de una imagen) que aporta calidad de definición y es útil para ampliaciones de tamaño de la imagen y para la visualización en pantallas grandes.

2. Su sensibilidad o rango dinámico (Dynamic Range), que dependerá fundamentalmente del área total del fotodiodo y del tamaño del sensor. El rango dinámico (latitud, sensibilidad, en ocasiones también llamado contraste o contraste máximo) es el valor entre el mínimo y el máximo nivel de la señal; es decir, la relación entre el valor de negro puro y el valor de blanco puro, lo que puede dar como resultado imágenes más realistas y detalladas.

En cuanto al tamaño del sensor, implica una regla sencilla: Si contamos con un sensor APS-C (por poner un ejemplo muy común en HD), sólo podremos aumentar la resolución (tener más megapíxeles) sacrificando la calidad de cada uno de ellos (pues resultarán fotodiodos con menor área sensible). Esta regla, como veremos, no es la única que determina la calidad de una cámara, pero sí resulta importante tenerla siempre presente.

4.8 Nivel de Sensibilidad/Ganancia del sensor (número o valor ISO):

El valor ISO en una cámara digital se refiere a la sensibilidad del sensor. ISO es una abreviatura de la Organización Internacional de Normalización, que establece estándares para la sensibilidad de la película fotográfica.

Las cámaras digitales permiten ajustar el valor ISO en una amplia gama, generalmente desde un valor bajo, como ISO 100 o 200, hasta valores más altos, como ISO 800, 1600, 3200 o más. Cuanto mayor sea el valor ISO la cámara podrá capturar imágenes más brillantes en condiciones de poca luz. No recoge más cantidad de luz, sino que aumenta la señal eléctrica que genera. Un valor ISO bajo, como ISO 100, es adecuado para condiciones de buena iluminación, como luz natural brillante o situaciones de estudio con iluminación controlada. Las imágenes tomadas con un ISO bajo tienden a tener menos ruido o granulado y colores más precisos.

En la imagen electrónica el ruido es aquella parte de la información que no se corresponde con la realidad que se quiere mostrar, originada por imperfecciones de los procesos de captación. En fotoquímico, el ruido toma el nombre de "grano", y se origina por una deficiente exposición de los elementos fotosensibles o un excesivo forzado en el relevado.

Un valor ISO alto, como ISO 800 o superior, se utiliza en situaciones de poca luz, como interiores con luz tenue o situaciones nocturnas. Es importante tener en cuenta que, si bien el ISO alto puede ser útil en condiciones de poca luz, también puede degradar la calidad de la imagen al introducir ruido. Por lo tanto, es recomendable usar el ISO más bajo posible que permita obtener la exposición deseada, y solo aumentar el ISO cuando sea necesario para evitar el uso excesivo de valores altos que puedan afectar negativamente la calidad de la imagen.

Explicado de otra manera, si el valor de carga acumulado por cada fotosito es demasiado pequeño, debe ser amplificado antes de su digitalización. El nivel de amplificación depende de la sensibilidad ISO que hayamos seleccionado en la cámara.

Al aumentar la sensibilidad ISO no estamos aumentando realmente la sensibilidad del sensor. Lo que estamos haciendo es amplificar la señal captada por el sensor antes de proceder a su codificación y registro; es decir que no estamos permitiendo una mayor entrada de luz, sino que estamos aumentando la amplitud o potencia de una señal electrónica.

Con el diafragma y la velocidad de obturación, el ISO supone la tercera forma de controlar la luz que vamos a registrar. Como ya vimos en una clase anterior, a esto se le llama de manera muy común el triángulo de exposición.

¿Qué son los sensores ISO-Less y cómo funcionan?

Los sensores ISO-less son capaces de controlar varios tipos de ruido, por lo tanto consiguen que nos dé lo mismo aumentar la ISO a la hora de disparar, que subexponer la toma y aumentar la exposición luego en el procesado, porque vamos a generar un ruido muy similar.

5 Resolución:

La resolución es un concepto sencillo: indica el número de píxeles de los que está formada una imagen digital. Cuando hablamos de resolución nos podemos referir en concreto a una resolución de un sensor, de una pantalla o de una imagen en general.

Un píxel o en plural píxeles (acrónimo del inglés picture element, 'elemento de imagen'), es la menor unidad homogénea que forma parte de una imagen digital. Cuanto mayor número de píxeles, mayor resolución, lo que significa mayor calidad.

En fotografía digital se suele hablar de la resolución en términos de "megapíxeles" o millones de píxeles. En cine, televisión y pantallas en general lo habitual es referirse a la resolución por una sola cifra: bien la del número de líneas o bien el número de columnas de píxeles. Pero el sentido es el mismo: contar los píxeles totales de una imagen.

En televisión, la resolución HD (High Definition) es el formato de vídeo estándar actual utilizado en la mayoría de emisiones: televisión digital terrestre (TDT), televisión por cable, televisión por satélite y vídeo por streaming, y establece entre otras cosas una resolución de 1920x1080. Aunque en muy poco tiempo la resolución 4k lo sustituirá.

1920x1080 supone que hay 1920 líneas de píxeles verticales y 1080 horizontales. Y en el cine y lo más común hoy en día es la resolución 4K, que correspondería a una resolución de 4096x2160 líneas de píxeles, mientras que el 4K para televisión está planteado con una resolución de 3840x2160.

5.1 Sistemas PAL y NTSC

PAL es la sigla de Phase Alternating Line (en español 'línea de fase alternada'). Es el nombre con el que se designa al sistema de codificación utilizado en la transmisión de señales de televisión analógica en color en gran parte del mundo, España incluida.

Cuando los 25 fps los tenemos divididos en 2 campos cada frame, es decir 50 campos por segundo en total, lo podremos ver como 50i. Veremos el tema de los escaneados y los campos más adelante.

En analógico había 625 líneas, pero nos quedamos con estas en digital para la velocidad de refresco.

Para que un formato sea considerado HD tiene que tener al menos el doble de cantidad líneas que el formato del NTSC

5.2 Relación de aspecto (aspect ratio)

La relación de aspecto o formato de pantalla es la relación entre el número de columnas y el número de filas de nuestra pantalla (entre el ancho y el largo, si se quiere). Se suele expresar bien por una fracción (por ejemplo 16/9 para la señal Hd de 1920x1080), o por el resultado de esta división (1,78) o también por la relación expresada en referencia a la unidad (1:1,78).

Además de hablar de relación de aspecto de una pantalla, nos podemos también referir a un sensor ó un píxel. La relación de aspecto del píxel no tiene nada que ver con la relación de aspecto de la imagen resultante o la relación de aspecto del sensor.

5.2.1 Relación de Aspecto del Píxel:

En cuanto a la relación de aspecto del píxel, en un momento de transición del sd al hd, se inventó el píxel no cuadrado para poder ver en 16:9 imágenes grabadas con un sensor de 4:3. Y para ello y para no cambiar toda la cadena de los sensores de las cámaras se anamorfizan los píxels (se estiran horizontalmente).

Después cuando emitimos hd con píxel no cuadrado ponemos un metadato al archivo para que se desanamorfe y no salga una imagen deformada.

En la resolución 4K (UHD: Ultra High Definition) ya es todo píxel cuadrado.

El cine no comparte ni tamaño de pantalla ni relación de aspecto con la televisión.

Después de todo, en una sala de cine nosotros podríamos proyectar casi cualquier formato de pantalla, simplemente recortando o “cacheando” (del inglés catch) la imagen, que sale por el proyector.

No obstante, y para evitar precisamente una excesiva variedad, la industria cinematográfica estableció en diferentes etapas de su historia algunas relaciones de aspecto consideradas estándares. La clásica fue en su tiempo la 1:1,37, denominada formato académico (Academy), y que fue la que adoptó casi igual (1,33 ó 4:3) la televisión en sus inicios. Una vez que la televisión adopta un formato, como sabemos (la cadena de hd), ya es difícil que pueda cambiarlo, pues implica cambiar toda la cadena de equipamiento. Sin embargo, en cine no supone mucho problema ni coste utilizar otras relaciones de aspecto, y así surgieron varios formatos muy panorámicos, generalmente intentando diferenciarse de lo que ofrecía su gran competidora, la televisión.

Los formatos muy panorámicos han pasado de moda y a día de hoy el cine suele trabajar con sólo dos relaciones de aspecto: 1:1,85 y 1:2,35.

Al tener diversas opciones, en cine se adoptó la costumbre, mucho antes de que se popularizara la alta definición, de nombrar los ficheros no por el número de líneas o píxeles verticales, sino por el número de píxeles horizontales. El número de píxeles horizontales siempre permanece inalterable, y el de los verticales vendría dado por la relación de aspecto que escojamos.

En la televisión la relación de aspecto pasó del anterior 4:3 al 16:9, sobre todo con la llegada del HD a partir de 2009, y así se pudo poder ver las películas de cine en la televisión a la misma o más parecida relación de aspecto.

Si hemos decidido, por ejemplo, que nuestra película para cine tenga una relación de aspecto de 1:1,85 de 4K, basta con hacer una simple operación para obtener una resolución vertical que significa $4 \times 1.024 = 4.096$ píxeles de resolución horizontal.

$4.096/1,85 = 2214$ píxeles (el redondeo es obligatorio, pues no podemos contar con decimales de píxel).

6 Conversión Analógico Digital A/D.

Conversión analógico a digital: es el proceso por el cuál se toman valores discretos (o muestras) de la amplitud un fenómeno continuo (es decir la señal de los valores recogidos por el sensor) para transformar la información analógica en información digital.

Frecuencia de muestreo, o Sampling rate: es la cantidad de muestras que se toman por unidad de tiempo. Si la velocidad con que se toman las muestras no es suficiente, no será posible reconstruir la señal original sin errores. El muestreo hace discreta la frecuencia de la onda.

Para entender el proceso, primero hay que entender cómo funciona el lenguaje digital, es decir el sistema binario. A diferencia del sistema decimal que emplea 10 dígitos del 0 al 9, el sistema binario emplea solo dos: el 0 y el 1, llamados bits. Esto viene condicionado por el funcionamiento de los componentes electrónicos que solamente admiten dos estados (Nivel alto “1” y el Nivel bajo “0”).

El teorema de muestreo de Nyquist: dice que el número de muestras ha de ser al menos del doble de la mayor frecuencia contenida en el espectro de la señal. Muestreando esta cantidad se puede representar correctamente la forma de la onda. Si no realizamos este número mínimo de muestras se genera el aliasing o el moaré que vimos en el apartado anterior. Es decir que en un caso donde las frecuencias son demasiado altas para las muestras que estamos recogiendo, se generan fallos.

Sin embargo, cuanto menos muestreo tengamos, menos espacio de discos ocupa la información, más fácil de manejar son nuestros datos en los ordenadores y más barata sale la producción.

6.1 El Muestreo de Color y el Submuestreo de Color

El muestreo en el contexto del audio digital:

Ya adelantamos que la frecuencia de muestreo estándar para audio de alta calidad es de 44.1 kHz, que significa que se toman 44,100 muestras por segundo.

Como el rango entre 20 a 20,000Hz forman los límites absolutos del rango de audición humana, la frecuencia de muestreo estándar para audio de alta calidad es de 44.1 kHz (el doble), que significa que se toman 44,100 muestras por segundo.

El muestreo de una señal de video es más complejo debido a que implica la captura de información visual en dos dimensiones (espacio y tiempo).

En la resolución espacial, cada cuadro o imagen de video se divide en una cuadrícula de píxeles, y cada píxel tiene valores de luz y de color que se registran (por ejemplo, 1920x1080 píxeles para video Full HD). Es decir, que el muestreo nos dirá cuántos de los píxeles son efectivamente contabilizados. El muestreo por lo tanto hace referencia a la cantidad de píxeles que se usan para obtener información de color, es decir, que se usan como muestra. Por lo tanto, cuanto mayor sea el muestreo de color, mayor cantidad de información habrá.

La frecuencia de muestreo de video en cuanto al parámetro del tiempo se expresa en términos de la cantidad de cuadros por segundo (Frame Rate, Cadencia, FPS, por sus siglas en inglés).

En la señal de video digital tenemos entonces estas dos referencias: el luma (luminancia o Y) y el croma (Crominancia o C).

La luminancia se compone con la siguiente proporción: $Y = 0.299R + 0.587G + 0.114B$, que se suele redondear para decir que la luminancia es la suma del 30 % de la señal del canal rojo, más el 60 % del verde, más el 11 % del azul.

La Crominancia o información tricolor para nuestro conocido sistema RGB es la cantidad de colores que vamos a poder registrar. Está formada a su vez por dos señales denominadas arbitrariamente U y V. La fórmula es:

$$C = U + V$$

$$U = B - Y \text{ (luminancia menos el componente rojo)}$$

$$V = R - Y \text{ (luminancia menos componente azul)}$$

Para poder transmitir esta información separada, se utiliza un proceso que se denomina codificación de señales en componentes (digitalización PCM ó YUV).

Una señal digital muestreada 4:4:4 tendrá siempre contabilizados todos los píxeles (con 3 canales RGB de muestras totales). El formato de muestreo 4:4:4 significa que no se realiza ningún submuestreo de croma, es decir, tanto la luminancia como la crominancia se muestrean en cada píxel. Esto resulta en una representación de color más precisa y detallada, ya que cada canal de color tiene la misma cantidad de muestras.

Es decir, que una señal digital muestreada en 4:4:4 mantiene la integridad total de la información de color, lo que es esencial en aplicaciones profesionales como la edición de video, efectos visuales y trabajos de posproducción donde la calidad del color es primordial.

El submuestreo de color en el video tiene sus raíces en la evolución histórica de la televisión. Para mantener la compatibilidad con los antiguos monitores en blanco y negro y para optimizar el uso del limitado espectro radioeléctrico, se decidió reducir la cantidad de información transmitida. Esto se logró eliminando parte de la información de color capturada por la cámara, haciendo la señal más manejable y eficiente en términos de ancho de banda.

En la televisión en B&W había una única señal de luminancia (Y), con intensidades ó voltajes intermedios que eran todos los grises. Con la llegada de la tv en color, en la misma señal había que transportar la señal del RGB. Esta codificación RGB traía problemas de estandarización entre sistemas PAL y NTSC en cuanto a parámetros de frecuencia de muestreo, cuantificación digital y codificación.

Como el ser humano es más sensible a la diferencia de brillo que a las diferencias de color, el submuestreo como medio para la compresión de la señal se basa en este principio: comprimir más el color que el brillo. Es decir, que para mantener una calidad aceptable de la imagen nos bastaría con tomar la mitad de muestras en cada señal de crominancia respecto de las que necesitamos para la señal de luminancia. La razón de comprimir una señal es para economizar ancho de banda y por lo tanto presupuesto. Al disminuir la resolución de las componentes de color (crominancia) y mantener la resolución completa de la luminancia, se puede reducir significativamente la cantidad de datos sin afectar perceptiblemente la calidad de imagen para el espectador promedio.

Así, la nomenclatura 4:2:2 definida en la Norma BT 601 de televisión digital significa que por cada cada 4 muestras de luminancia se toman sólo dos muestras de la información de color (2 de cada 4 píxeles).

Existen también codecs de inferior calidad con un tipo de muestreo de 4:2:0 o 4:1:1. El 4:2:0, que quiere decir que se tomará en una primera línea de barrido 4:2:0 y en la siguiente línea 4:0:2 y así sucesivamente hasta completar el barrido completo.

Si pensamos en una señal HD 1.080, sabemos que cada canal RGB necesitará 1.920 x 1.080 píxeles o muestras para un muestreo completo 4:4:4. Habrá pues $1.920 \times 1.080 = 2.073.600$ muestras del canal rojo, otras tantas del canal verde y otras tantas del canal azul.

Sin embargo, en 4:2:2 tendremos 1.920 x 1.080 muestras de luminancia, y sólo 960 x 1.080 de la señal U y otras 960 x 1.080 muestras o píxeles activos del canal V. A pesar de que visualmente es apenas apreciable, este hecho puede producir distorsiones y falta de precisión en procesos de postproducción. Un caso típico es el uso de un chroma key o incrustación por llave de color, que siempre será más preciso, con bordes más definidos, trabajando en 4:4:4.

6.2 Sobremuestreo

En oposición al submuestreo, existe la técnica conocida como sobremuestreo (super-sampling, binning). Consiste básicamente en que un píxel en la imagen toma la información no de una única celdilla o fotodiodo, sino de dos o más.

El objetivo es reducir el aliasing y, mucho más interesante, obtener imágenes de mayor rango dinámico (Higher Dinamic Range images).

El sobremuestreo es una de las técnicas de mayor futuro y recorrido en los próximos años.

6.3 Efecto Moiré o aliasing y Filtro anti-aliasing (AA, OLBF)

El efecto moiré es el resultado de patrones que se superponen o intersecan en una imagen, mientras que el aliasing se relaciona con el submuestreo o muestreo inadecuado de señales analógicas o imágenes, lo que conduce a artefactos visuales no deseados en la representación digital. Cuando se captura o se visualiza una imagen o un video con una resolución insuficiente o una tasa de muestreo inadecuada, se pueden producir efectos de aliasing en frecuencias espaciales altas como ropa como en la foto del ejemplo. En las cámaras, para evitar este defecto, se usan medios físicos, como el filtro AA (anti-aliasing), también conocido como de paso bajo (Optical Low Band Filtei OLBF) o por software.

El filtro AA suele consistir en una fina capa de un material determinado que recubre el sensor. El material en concreto varía y suele ser parte del secreto industrial del fabricante. Lo que hace es discriminar las "altas frecuencias de contraste", líneas finas muy definidas; de ahí el nombre de paso bajo. La introducción de un OLBF significa, inevitablemente, una pérdida siquiera mínima en la "nitidez" (entendida como líneas por milímetro).

Marco futuro:

Tecnológicamente, hoy no supondría ningún problema que toda la industria televisiva trabaje ya en RGB. Excepto en el caso de las emisiones, por el tema ya comentado de la limitación del espacio radioeléctrico. La necesaria compatibilidad con los equipamientos previos, permiten que el 4:2:2 siga siendo el estándar televisivo profesional, como recogen todas las normativas y recomendaciones internacionales. No es impensable suponer que a medio plazo las opciones de grabación y postproducción de todas las cámaras sean 4:4:4. El 4:2:2 se mantendrá como una conversión final para su posible emisión por las televisiones y canales tradicionales.

6.4 Profundidad de color, profundidad de bits, cuantificación, cuantización o Sampling depth.

Nos podemos encontrar con cualquiera de estos nombres.

Cuantificar es un verbo que se refiere a la acción de enunciar una cantidad, en nuestro caso consiste en dar una cantidad de bits a un píxel para darle profundidad (en el sentido de calidad). Para eso, en la cuantificación se asigna un número de colores/valores a cada canal RGB para

dentro de nuestras posibilidades técnicas dar toda la calidad posible. Esto es dicho con otras palabras la cantidad de matices de luz y color que podremos obtener, y por lo tanto con cuánta exactitud vamos a reproducir los colores de la realidad.

Para entender el proceso, primero hay que entender cómo funciona el lenguaje digital, es decir el sistema binario. A diferencia del sistema decimal que emplea 10 dígitos del 0 al 9, el sistema binario emplea solo dos: el 0 y el 1, llamados bits. Esto viene condicionado por el funcionamiento de los componentes electrónicos que solamente admiten dos estados (Nivel alto “1” y el Nivel bajo “0”).

Como decimos, la manera es asignando un número de bits para codificar y guardar la información de color de cada píxel en una imagen. Cuanto mayor sea la profundidad de color en bits, la imagen dispondrá de una paleta de colores más amplia.

Recordamos que un bit es una cantidad de memoria que puede tener el valor 0 ó 1 y por lo tanto una imagen de 1-bit sólo podría representar dos valores, es decir, una imagen en blanco ó negro, sin grises. (0=color negro, 1= color blanco).

2-bits= $2^2=2 \times 2=4$ colores (00=color negro, 01=color X, 10=color Y, 11=color blanco).

3-bits= $2^3=2 \times 2 \times 2=8$ colores.

4-bits= $2^4=16$ colores

Y así consecutivamente, por lo tanto cuando hablamos por ejemplo de un modelo en uso, el de 8 bits (en realidad 8 es el exponente):

$2^8=256$ niveles o tonos de gama

Se suele hablar de bits por canal, y no del total de bits, pues el número de canales puede variar. Una imagen en blanco y negro, por ejemplo, sólo tiene un canal, el de la luminancia.

Como hemos dicho, en color contamos con tres canales, RGB, por lo que si calculamos el número de combinaciones posibles de colores en una señal de 8 bits por canal, tenemos que $256 \times 256 \times 256 = 16.777.216$ diferentes tonos de colores representables en cada píxel.

O puede haber una imagen con los tres canales (RGB) más un cuarto, el “Canal alfa”, que permite albergar una información de transparencia.

6.4.1 Canal Alfa:

En computación gráfica, la composición alfa o canal alfa es la que define la opacidad de un píxel en una imagen. El canal alfa actúa como una máscara de transparencia que permite, de forma virtual, componer (mezclar capas) imágenes o fondos opacos con imágenes con un cierto grado de transparencia.

Así, en un formato de imagen que admita transparencia, a los tres canales de color básicos RGB que definen la cantidad de rojo, verde y azul respectivamente del píxel, se añade un cuarto canal, el alfa, que define el grado de opacidad de ese píxel.

La presencia del canal alfa es una de las ventajas del formato de imagen PNG frente a otros como el JPEG, el cual no posee información alfa, y, por tanto, no admite transparencias. Antes de la creación de este canal, y cuando no había ordenadores, en fotografía y en el cine se empleaban máscaras físicas (recortadas o pintadas) que tapaban las partes de la imagen que no se debían exponer a la luz y que servían para generar una segunda exposición controlada que permitiera componer dentro de la película elementos que no estaban presentes en la primera exposición o

rodaje. También se utilizaba habitualmente la técnica de pintar imágenes sobre fondos transparentes, como acetatos o vidrios. Con la creación de sistemas informáticos, el canal alfa resolvió la necesidad de recrear esta transparencia que permitiera fotocomponer varias capas.

La cuantificación funciona tomando valores de amplitud de señal analógica y después cuantifica con bits dichos valores de amplitud mediante niveles. Ojo, el muestreo hace discreta la frecuencia de la onda y la cuantificación hace discreta la amplitud de la onda.

Este proceso supone que con más bits obtendremos más información sobre la amplitud, o también mejor dicho mayor precisión en la información de la amplitud.

La diferencia entre el valor real de la muestra y su aproximación a un número entero se denomina error de cuantificación.

Como dato tened en cuenta que el almacenamiento se expresa en Bytes. Y la velocidad de transmisión de la información digital en bits. 8 bits=1 Byte

6.4.2 El estándar de 8 bits.

Aunque en la actualidad la tendencia es utilizar 10 bits por muestra de forma casi universal, el grado de compromiso para el mundo de la HDTV se fijó en 8 bits por canal.

6.4.3 Banding:

El banding aparece cuando la tecnología de la que disponemos, no puede generar suficientes tonos para representar un degradado de color tan rico. Normalmente ocurre a causa de la compresión de la imagen, es decir, nuestro archivo tiene un límite en su información o los programas de edición tienen también un límite en la representación de los colores.

Por eso, cuando no hay suficiente información o tonos para representar ese cambio tan gradual en un degradado, aparecen las líneas del banding. Estas líneas no son ni más ni menos que ese salto entre un tono y el siguiente cuando no hay suficientes colores para que se vea una transición gradual.

6.5 Gamuts:

Los Gamuts son los espacios de color que vamos a utilizar para representar o captar en cámara todo el rango de colores. Nos van a indicar en qué nivel está la señal RGB, el punto de blanco y la cantidad de saturación máxima. Es decir, no indica cuántos colores, sino que básicamente no es más que un eje de coordenadas para explicar los niveles de saturación de cada color.

Es la cantidad del espectro electromagnético visible para los humanos al que puede dar cobertura la tecnología y los dispositivos, tanto en la captación, la codificación, el tratamiento y la transmisión de imágenes.

La tecnología de las cámaras puede grabar en gamuts más altos de lo que luego pueden reproducir los dispositivos. Este es uno de los grandes problemas de los monitores, que no llegan a representar lo que sí se puede captar con las cámaras de última generación, con gamuts mucho más amplios. Hay algunos incluso que se salen del espectro visible. Y hay que tener en cuenta que aunque grabemos con estas cámaras a la hora de retransmitir la señal por ejemplo a un monitor de tv, se adoptará la norma 709. Es decir, que existen muchos Gamuts diferentes, cada modelo de cámara te trae uno prácticamente, y al igual que en otros procesos, hay que tener en cuenta que

para optimizar los recursos, la captación y codificación en una cámara debe llevar el mismo tipo gamut que luego la representación en pantalla.

Los monitores profesionales de grado 1 tienen capacidad para las características de la norma 2020 u otros sistemas de color como el XYZ para el cine, que no lo hemos dado.

6.6 Rango Legal y Rango Extendido

Una cuestión importante de los archivos de video es el nivel máximo y mínimo de brillo permitido. Las normas técnicas de difusión televisiva establecen unos niveles determinados para el blanco puro y el negro puro, pero en exhibición de cine digital o en las pantallas ordenador se utiliza una norma distinta. Por lo tanto es necesario configurar los niveles de brillo y contraste del master de nuestro proyecto en función del uso que se les vaya a dar: cine, internet o televisión. Se denomina **rango legal** a los niveles de señal para televisión. El material ajustado para cine digital utiliza el **rango extendido**, al igual que las pantallas de ordenador. Por ejemplo para una codificación a 8 bits de la señal de luminancia se disponen de 256 valores que se corresponden con una gama de grises. En el rango legal el valor máximo permitido que se usa para el blanco puro es de 235, y el negro puro se sitúa en el valor 16. Cuando se dispone de 10 bits para poder codificar la señal, el rango total de valores disponibles es de 1024 puntos, en este caso los valores del rango legal están entre el 940 para el blanco y el 64 para el negro.

Un ejemplo: Si tenemos un nivel de negros por ejemplo al 16 porque lo queremos en rango legal y nosotros lo estamos viendo gris, es porque estamos interpretando mal alguno de los procesos, puede ser un error en la grabación o una monitorización deficiente. Por esto es tan importante el monitor de forma de onda.

Siempre todos nuestros archivos de video van a trabajar en modo legal o modo extendido, uno u otro, y es importante que tengamos claro cómo trabajamos en toda la cadena desde la grabación, y edición para luego entregar a la emisión. Suele coincidir que los codecs con cuantificación 4:4:4 vayan asociados al rango extendido y los 4:2:2 vayan con rango legal.

Los monitores profesionales tienen la opción de monitorizar en legal o en extendido.

Siempre que trabajemos en un monitor de ordenador normal estaremos viendo en extendido pero nosotros podemos colocar nuestros niveles en rango legal para en este caso exportar un proyecto para la televisión.

En la UHD TV (4K) se sigue utilizando el rango legal.

Una última forma de asimilar este tema para ver las cosas desde diferentes puntos de vista:

Rango Legal y Rango Extendido correctamente interpretados se ven correctamente, pero si es necesario saltar de uno a otro hay que modificar los parámetros de brillo y contraste.

Ejemplo de codec en rango extendido pero interpretado como legal por el software. Se ve con menos contraste, aparece con más brillo, una imagen que suele decirse más lavada.

Y al revés, un rango legal interpretado como extendido se vería con más contraste porque los valores se estiran.

7 Sistemas de Escaneado: “P” (progresivo) e “I” (interlaced ó entrelazado). Y sistema mixto PSF.

Un sistema de escaneado en televisión se refiere al método utilizado para captar, actualizar y mostrar imágenes en una pantalla mediante la exploración de líneas y cuadros en una secuencia específica. Este aspecto técnico hace referencia al método de exploración secuencial de las líneas de una imagen de televisión. Así, existen dos sistemas de escaneado o barrido:

Entrelazado: El origen del barrido entrelazado se debe a dos motivos: por un lado, evitar

el efecto de excesivo parpadeo que tienen las cadencias bajas (24, 25 o 30 ips lo son). Y por otro lado, corregir un defecto técnico de los televisores de tubos catódicos que están en el origen de la televisión y reducir el ancho de banda.

La solución fue hacer un barrido entrelazado de la imagen. En este sistema, una imagen completa o cuadro o frame, se forma dividiendo cada cuadro en dos campos (fields) diferentes: el primero escanea todas las líneas impares, y el segundo escanea las líneas pares. Estos campos se entrelazan para formar un cuadro completo. Este método es comúnmente denotado por "i" en la descripción de la resolución de la pantalla, como 1080i. También nos lo podemos encontrar escrito como 25i (frames por segundo) ó como 50i (que significa lo mismo: 25 frames por segundo en 2 campos separados).

Así pues, una imagen de 1.080 i está compuesta de dos campos diferentes, cada uno de ellos de 540 líneas. Uno corresponderá a las líneas impares o superiores (even, up): 1, 3, 5, 7, 9.....hasta 1.079; y otro a las líneas pares o inferiores (odd, down): 2, 4, 6, ..., hasta 1.080.

En todos los codecs de HD se coloca primero el campo impar.

Progresivo: En este sistema, la imagen se forma línea por línea de manera secuencial, de arriba a abajo. Cada cuadro completo se escanea en una sola pasada, y se muestra de manera progresiva en la pantalla. Es comúnmente denotado por "p" en la descripción de la resolución de la pantalla, como 1080p. Este método es utilizado en tecnologías como los televisores de alta definición (HDTV) y monitores de ordenador. Hoy día, las pantallas de televisión a partir de resoluciones 4K no utilizan el barrido entrelazado, sólo progresivo. Pero la emisión televisiva se hace todavía en entrelazado.

1080i y 1080p muestran una resolución de 1920×1080, lo que ocurre es que en “p” actualiza todos los puntos de la imagen 25 veces por segundo, mientras que en “i” trabaja igual a 50 Hz pero representa primero unas líneas y luego otras, es decir, se reparten los 50 Hz: 25 Hz para las pares y 25 Hz para las impares.

Hay que tener en cuenta que una grabación entrelazada exige una emisión entrelazada y que en caso de duda es mejor captar la imagen en progresivo cuya conversión a entrelazado no ofrece ningún problema. En cambio, si grabamos en entrelazado y lo pasamos a progresivo tendríamos el llamado efecto diente de sierra.

Existe un tercer formato de tipo mixto:

PSF: Progressive Segmented Frame: El fotograma segmentado progresivo es un esquema diseñado para trabajar con vídeo de barrido progresivo utilizando equipos entrelazados. Es decir una captación de cámara progresiva y un tratamiento o guardado entrelazado. Con Psf un cuadro progresivo se divide en dos segmentos con las líneas impares en un segmento y las pares a otro. Es

decir, que se genera una señal en P desde la cámara y luego cuando llega a la emisión, a un grabador o a un mezclador, hay un proceso en el que se divide la señal creando un “falso” entrelazado.

Se considera progresivo a efectos estéticos.

En los sistemas de definición estándar, el PAL o NTSC se usa el (i), en HD (p ó i) y UHD sólo P.

7.1 Ventajas y desventajas de cada barrido: resolución dinámica

Por lo general, se entiende que los sistemas progresivos tienen una mayor calidad debido a que su resolución real es el doble que el mismo formato en entrelazado. Sin embargo, no podemos olvidar que el entrelazado tiene más “resolución dinámica”. Toma el doble de muestras de la realidad por segundo que el progresivo. Produce sensaciones más fluidas, con menor efecto parpadeo.

Por ejemplo en una retransmisión de atletismo de una carrera de 100m (unos diez segundos grabados a 25 frames por segundo), el progresivo captará 250 imágenes de 1.080 líneas, pero el entrelazado tendrá 500 imágenes de 540 líneas. ¿Quién ofrecerá mejor reproducción de la realidad de cara al espectador? En este caso, probablemente el entrelazado. La carrera tendrá una representación más fluida en entrelazado que en progresivo.

Ventajas I: Nitidez, movimientos de cámara, títulos de crédito rápidos.

Desventajas I: Tienes que configurar la dominancia de los campos y funciona mal en dispositivos P. Puede generar posible parpadeo en imágenes congeladas.

Ventajas P: Funciona bien en todos los dispositivos, no tienes campos.

Desventajas P: Limitado en la velocidad de los movimientos de cámara y no permite los títulos rápidos.

En cuanto a las diferencias estéticas, el barrido “i” se suele decir que tiene una estética más de video y el “p” un estilo más cinematográfico.

7.2 Cadencias del futuro

La solución sencilla para disfrutar de todas las ventajas del progresivo (mejor resolución horizontal) y las del entrelazado (mejor resolución dinámica) sería doblar las cadencias estándares.

Como ya sabemos, los tradicionales 24 fps del cine son realmente pocos, lo cual obligaría a que en las salas se proyecte dos veces cada fotograma para evitar el efecto parpadeo. Por eso se propone pasar a cadencias digitales de 48 fps nativas.

8 Visualización de la señal

Uno de los problemas más comunes en el trabajo digital es la correcta monitorización de la imagen. Sólo los monitores perfectamente calibrados pueden asegurarnos un flujo de trabajo correcto. Pero esta opción no siempre es posible, sobre todo durante un rodaje.

Por descontado, los visores propios de las cámaras, dado su pequeño tamaño y construcción, sólo son fiables para el encuadre y no para ofrecer precisión de calidad. Para un control más preciso de la señal digital se utilizan los visores (scopes) de la información propiamente dicha. Los más comunes son:

El Monitor Forma de Ondas, el Vectorscopio y el Histograma.

Los dos primeros son usuales en la televisión, tanto analógica como digital, y el último en el software de postproducción y cámaras.

8.1 El Monitor Forma de ondas

La forma de ondas (waveform) nos da información, fundamentalmente, de la intensidad de la señal. El eje horizontal corresponde a las líneas de resolución y el vertical propiamente a la intensidad de cada línea. Ya sabemos que si hablamos de una señal de ocho bits, por ejemplo, la base corresponde a 0 y el punto máximo a 255. Y que en rango legal, o broadcast safe o señal televisiva estándar se sitúa entre 16 y 235.

Podemos seleccionar independientemente la luminancia o la crominancia (en el caso de señales YUV) o los canales rojo, verde y azul (en el caso de señales RGB); o varios de ellos a la vez.

8.2 Vectorscopio

El vectorscopio (vector scope) nos dará información tanto del tono de color dominante como de la saturación.

Se presenta como un círculo que corresponde a la representación del círculo cromático que ya conocemos, con los colores primarios (rojo, verde y azul) y sus complementarios (cian, magenta y amarillo) situados de manera opuesta.

8.3 Histograma

El histograma es tradicional en el mundo de la fotografía y muchas cámaras de vídeo y cine digital, nos da información sobre la cantidad de píxeles en cada una de las intensidades de luma y croma.

Una señal "rica" en el histograma tendrá variedad en todos los valores, pero principalmente en los medios. Generalmente (pero no siempre, pues todo depende de la imagen en concreto), en este caso adoptaría forma de campana.

El Histograma es una herramienta orientativa que sirve para ver si tu imagen está bien expuesta o no, pero no sólo eso, también puede ayudarte a ver si tu imagen tiene un rango dinámico rico o pobre, y por lo tanto, puede ayudarte a mejorarlo.

9 COMPRESIÓN

Dado el enorme flujo de datos que implica el video de alta definición y el cine digital, las empresas comerciales desarrollaron sistemas que comprimen los datos generados por la cámara para su grabación. Esta compresión estaba basada en la utilización de algoritmos matemáticos que conocemos como **códecs** (codificador-descodificador). Estos códecs utilizan diversos métodos para reducir la información de imagen a un flujo de datos más manejable, pero sin perder la "calidad subjetiva" de la misma. Entendemos por "calidad subjetiva" el hecho de que el espectador final no aprecie diferencias significativas que estorben el correcto entendimiento de los datos. Los códecs son más eficientes cuanto más comprimen la señal sin pérdida aparente de la calidad.

La compresión consiste, en esencia, en la simplificación o la eliminación de un tipo de información existente en todo mensaje: la redundante. En cualquier mensaje hay que diferenciar entre informaciones entrópicas y redundantes. La información entrópica es la mínima esencial para la decodificación de dicho mensaje sin error ni ambigüedad, y la redundante es aquella que no aporta nada nuevo ni varía la esencia de dicho mensaje. La cuantía de la reducción de la cantidad de datos que componen el mensaje suele expresarse en forma de relación numérica llamada tasa o ratio de compresión. Una tasa de compresión de 2:1 significa una reducción de la información del 50%, mientras que una ratio de 10:1 supone una compresión a una décima parte de la información original, esto es, una reducción del 90%.

Factores fundamentales que se definen dentro de un códec:

- 1- Compresión
- 2- Muestreo de color
- 3- Profundidad de bits
- 4- Resolución
- 5- Tasa de transferencia

9.1 Tipos de compresión:

Intraframe: La compresión se realiza dentro de cada fotograma. Coge un píxel de muestra y actúa sobre los de alrededor.

Interframe: Se utilizan varios fotogramas a la vez para realizar la compresión y solamente se guardarán las diferencias entre unos y otros.

Para este tipo de compresión utilizamos las llamadas Secuencias GOP (Group of Pictures).

I=Index. Es el fotograma original, completo sin compresión.

BB= Predicción Bidireccional: guarda sólo las diferencias que tenga el fotograma actual con el fotograma I.

P=Predecido: Una estimación partiendo del fotograma I y la evolución de los fotogramas intermedios de B.

Es decir, que la secuencia se fracciona en grupos de imágenes (frames) correlativas (GOPs. Group of Pictures) en número suficiente para aplicar la compresión temporal. Cada GOP cuenta con una primera imagen I (índice) y un número variable de predictivas P y bidireccionales B.

9.2 Ratios de compresión:

1:1, 2:1, hasta 22:1. Cada vez se consiguen ratios más altos sin pérdida de calidad.

9.3 Códecs: Media y Metadatos.

Los códecs son soluciones industriales que suelen tener una licencia o patente comercial detrás. Cuando esto es así, forman parte del secreto tecnológico de las empresas y se consideran por tanto como "códecs propietarios" o cerrados.

Necesitaremos tener licencia tanto para la codificación como para la decodificación o el simple visionado en nuestro ordenador de un archivo.

Cuando grabamos vídeo casi siempre lo hacemos con determinada compresión, es decir generamos un archivo con las características que nos interesen para un proyecto concreto. Aunque también se puede grabar sin compresión por supuesto.

Hay determinados codecs vinculados a determinadas cápsulas pero en general son independientes.

Los MXF OP Atom (Operation Patterns) son una buena opción un streaming de video para grabación, entre otras cosas no hace falta que esté cerrado el archivo para poder abrirlo e ir visionándolo, y también porque transporta muchos metadatos y audios por separado.

Tres de los codecs más usados son: H264, Pro Res y DNX.

El nuevo salto en este entorno de códecs es el MPEG4 o también lo encontramos como MP4 (H.265), más conocido como HEVC (High Enhanced Video Coding), el códec que se está utilizando para la Ultra Alta Definición (UHD). Este nuevo método de codificación de vídeo, que reduce a la mitad el ancho de banda, hace posible que se puedan ver contenidos en streaming con resoluciones Full HD o 1080p, no solo en televisión, sino también en otras plataformas como móviles o tabletas que serán los grandes beneficiados.

