

Capturando la señal de vídeo

Para comprender cómo funciona todo el mecanismo de grabación desde que encuadramos hasta que le damos al botón de grabación, la señal de vídeo pasa por diversos y complejos procesos. El sensor, los circuitos internos o el sistema de grabación empleado en una u otra cámara son fundamentales para entender las diferencias de por qué la luz es capturada y posteriormente grabada dentro de una cámara de HD.

Tomaremos como referencia propuestas tecnológicas basadas en sensores CMOS y captura RAW, así como apuestas inherentes a sensores con tres CCD.

Encontramos varias etapas entre la captura de datos y la grabación de la señal de video:

1. Formación de la imagen
2. Captura de la imagen en el sensor
3. Procesado de la señal (Sistemas HD / Raw)
4. Conversión A/D
5. DSP
6. Compresión
7. Debayering

1. Formación de la imagen

El viaje de la luz comienza al encuadrar el motivo con una lente, ya sea fija o de focal variable. Dentro de este motivo (paisaje, retrato...) existen diferentes valores tonales de gris y color, sobre la base de la propiedad de los objetos a reflejar una determinada longitud de onda.

La óptica es la encargada de que se forme una imagen lo más nítida posible. Además de marcar el valor de plano (de angular a tele) y la profundidad de campo, ayuda a determinar la luminosidad de la escena mediante el ajuste del diafragma.

Gracias a la propiedad de refracción de las lentes, la luz que reflejan los objetos dentro del motivo es transmitida hacia delante y focalizada en un sensor.

2. Captura de la imagen en el sensor

El siguiente paso del "viaje" de la luz en su captación y posterior grabación es el sensor. En este "chip" es donde tiene lugar la auténtica captación. La luz que forma parte de la escena o motivo -es decir, la realidad- pasa a través de la lente y alcanza una serie de elementos cargados eléctricamente llamados píxeles. Estos tienen la propiedad de reaccionar eléctricamente con mayor o menor intensidad

al punto origen al cual están asociados, es decir, al brillo de la imagen original. Para entenderlo mejor, supongamos que estamos encuadrando a una persona de noche con un proyector directo a su camisa blanca. Si tomamos una exposición "correcta", su piel está expuesta al 0,35v, su camisa a 0,67v y el fondo a 0,20mv (noche). Los píxeles que forman parte de la camisa reaccionaran con una mayor carga eléctrica que los que forman parte del fondo, donde prácticamente no habrá reacción y serán interpretados como ausencia de luz. Como veremos más adelante, dentro de un sistema digital, podríamos hablar de que el máximo brillo de luz es 1 y el mínimo, es decir, ausencia de luz (negro) es 0. Por supuesto, la variación de luz mediante el diafragma y el control de iluminación, afectará a todo este proceso de captación de la imagen.

Antes de que esta "energía lumínica", llegue a los elementos activos, ha de pasar por un separador cromático, que como si de un tamiz se tratase, dejará pasar una determinada longitud de onda. Es en este punto donde se puede establecer una de las principales divisiones tecnológicas en cuanto a sensores se refiere, por un lado encontramos el prisma dicroico, típico de los CCD, y por el otro, la denominada "máscara Bayer".

2.1 El prisma dicroico

Tanto los CCDs como los CMOS son monocromos, por eso la principal función de los separadores cromáticos es la de separar la luz que llega a través de la lente, en las tres longitudes de onda correspondientes a los colores primarios: Rojo, Azul y Verde. El prisma dicroico, muy utilizado en sensores CCD, consiste en una serie de espejos dentro del "bloque dicroico" que reflejan y dejan pasar las determinadas longitudes. Cada una de estas imágenes es enviada a cada uno de los 3 CCDs a los que están asociados, formándose 3 imágenes: R, G y B.

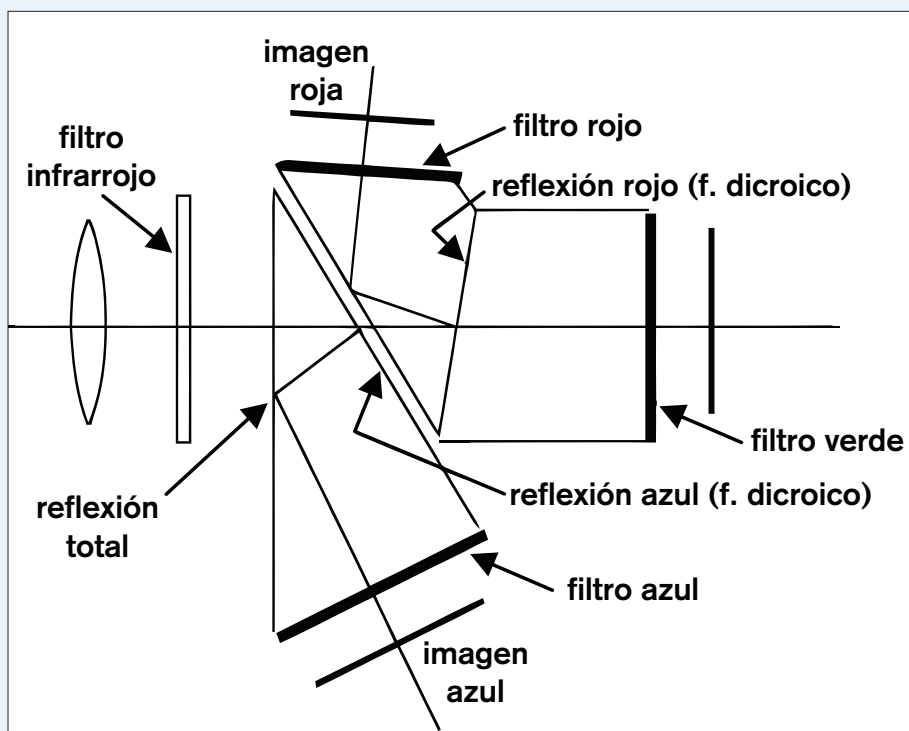


Fig.1 Prisma por dentro.

Este bloque permanece sellado para que no entre polvo, y una de las desventajas de utilizar este sistema es que dicho bloque sufre variaciones de temperatura que afectan al plano focal de la imagen, pudiéndose corregir mediante el aro de Back focus, que hace coincidir éste con los CCDs.

2.2 La máscara Bayer

Para separar el color se aplican filtros de color verde, rojo o azul sobre cada píxel, de forma discontinua encima del sensor. La "máscara Bayer" está muy relacionada con los CMOS, los cuales veremos, en el caso que nos interesa, desde el punto de vista de la captura RAW. Como en el caso anterior, esta máscara con forma de mosaico permite pasar una serie de longitudes de onda a cada uno de los píxeles. Como veremos más adelante, la extracción de la información de color tiene lugar durante el proceso de debayering o demosaicing (que viene a ser lo mismo).

2.3 CCD y CMOS

Un sensor puede ser básicamente de dos tipos: CMOS (Semiconductor complementario de metal óxido) y CCD (dispositivo de carga acoplada). Los píxeles constituyen la estructura básica del sensor, y en función del número integrado en su interior, tendremos más o menos resolución. Desde cámaras con resoluciones típicas de 600.000 (SD), pasando por los 2 millones del HD, hasta cámaras como la Red One (+4K - ultra HD). En la lectura de la imagen original se utiliza la exploración progresiva (25p) o entrelazada (50i).

En el modo entrelazado se capturan 25 fps de la siguiente forma: Primero se leen las líneas impares del cuadro durante 1/50 y posteriormente las pares durante el mismo periodo de tiempo, dando como resultado 1/25 por segundo el cuadro completo. Es decir, que un campo se captura en instantes de tiempo diferente.

De esta forma se consigue reducir el ancho de banda y que la imagen no parpadee, como ocurre con la exploración progresiva. Aquí, el cuadro se captura entero (todas las líneas) durante el mismo instante de tiempo, es decir, 1/25. Es exactamente lo mismo que hace una cámara de cine. El exceso de "fillage" que observamos es debido al uso de monitores de tubo (que trabajan en entrelazado) y al hecho de que hay mucha más profundidad de campo en HD, lo que dificulta el centrar la atención en el personaje.

En los CCD, la transferencia de salida de los píxeles (ya cargados) desde su localización física hasta la siguiente etapa de preamplificación puede ser de varios tipos (IT,

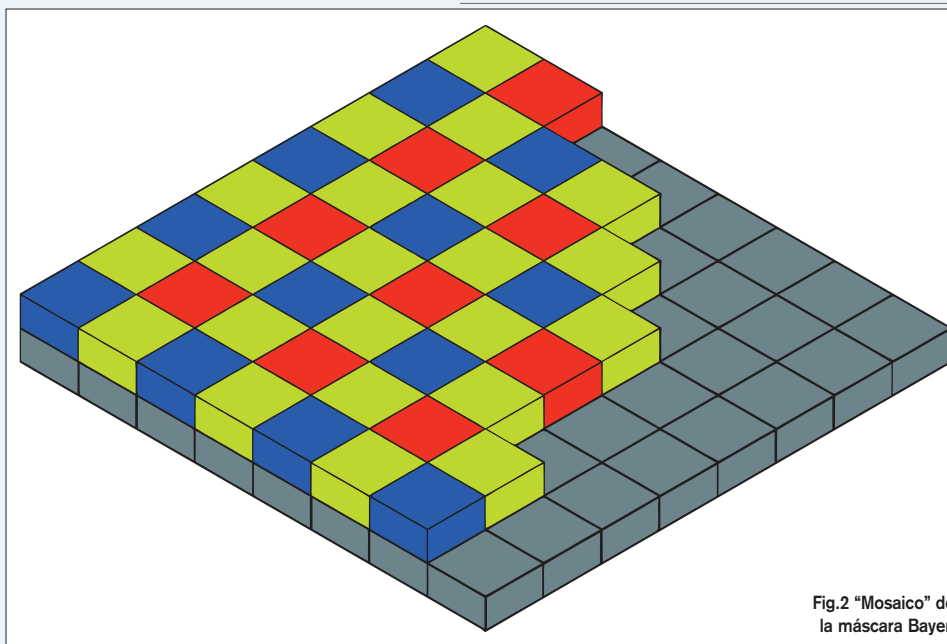


Fig.2 "Mosaico" de la máscara Bayer.

FIT...). Esto explicado muy básicamente sería la acción de captura-almacenaje-transferencia del elemento sensor.

Antes de pasar al siguiente proceso, algunas de las ventajas e inconvenientes relativos al diseño de CMOS y CCD:

- ❖ Al diseñar un sensor como único elemento fotoconversor de una cámara, se puede conseguir un mayor tamaño, compatible con ópticas de 35mm (y su importante profundidad de campo. Este es el caso de los CMOS y la máscara Bayer.
- ❖ Los CCD transportan la señal hasta los extremos para amplificarla y transmitirla hacia el conversor A/D, mientras en el CMOS cada uno de los píxel lleva incorporado un amplificador de la señal eléctrica y el conversor analógico-digital forma parte de su propia estructura. Esto provoca que el rendimiento de consumo de los CMOS sea menor, lo que representa una ventaja en la autonomía de la cámara.
- ❖ Al ser semiconductores no se requieren complejos procesos de fabricación siendo más económicos que los CCD.
- ❖ Quizá uno de los principales perjuicios era la presencia de mayor ruido en los CMOS que en los CCD, pero esto se está revirtiendo con la mejora de la tecnología actual. En contrapartida, y asociado a las cámaras con un único sensor mayor, la captación y posterior gestión de los datos son realizados por los CMOS con más agilidad que en los CCD, además de reducir el ruido.

3. Procesado de la señal

Es en este punto donde podemos establecer las dos grandes diferencias en cuanto al procesado de la señal se refiere. En los procesos con cámaras HD convencionales, la

señal es transformada y su origen queda "inservible", mientras que en los procesos RAW, siempre estamos a tiempo de recuperar la señal original.

3.1 Proceso típico señal cámara con procesamientos digitales

Antes de que la señal se digitalice (recordemos que nuestra señal actual es analógica), se establece un punto de **No retorno**. En otras palabras, estamos haciendo un "positivo" de una señal negativa pura. Como ventaja principal tiene la inmediatez, el poder utilizar la imagen en, por ejemplo, directos (noticias, deportes, eventos...) o en muchos casos ni etalonarla. Por supuesto, esto provoca una pérdida de calidad de la señal RGB original. La señal RGB que proviene de los sensores ha de ser convenientemente adecuada para su posterior grabación, emisión..., por lo que es introducida en una serie de PRE-amplificadores cuya función es la de estabilizar la señal, con mayor ganancia y menor ruido. En sucesivos procesos, la señal es convenientemente estabilizada gracias a lo que se conoce como procesos de cámaras, siendo más comunes:

- ❖ **Corrección de gama:** Compensa la falta de linealidad del TRC en el TV. Para ello se aplica una curva 0,45 (la normal en cualquier cámara de HD) que compensará la producida por un TV (coeficiente 2) dando como resultado una respuesta lineal (1). Este es uno de los motivos por el cual el "vídeo" / "HD" tiene tan mala fama en cuanto a latitud se refiere, ya que se sacrifica su rango dinámico para que la imagen sea reproducida correctamente en un TV. De hecho, sensores como por ejemplo el antiguo HAD ya llegaban a los 80db. La mayoría de sensores tienen una respuesta lineal a la luz,

pudiendo variar esta curva en lo que conocemos como curva de gamma de vídeo (0,45) o en el caso de cámaras como Viper, Genesis, F23, Red One, transformarlas a Log o Lineal (8, 10, 12, 14 bits) en el conversor A/D. Estas son más extremas en su rendimiento. Cámaras como la Varicam y la Sony F900, mediante el uso de Film Rec e Hypergammas respectivamente, no compensan el TRC por lo que se incrementa el rango dinámico. Esto conlleva que la imagen se vea más plana en un televisor, siendo necesario (o no) ajustarlas posteriormente. Las curvas Log o no lineales de la F23, RED, SI 2K, etc. van por el mismo camino, aunque de una manera más

efectiva, siendo posible recuperarlas mediante Luts para su correcto visionado en diferentes espacios de color.

- ❖ **Corrección de sombras:** Se corrigen errores típicos de sombreados introducidos por la lente o por la falta de uniformidad en el sensor. Posteriormente, el usuario tiene la posibilidad de "tocar" los menús de cámara para compensar este tipo de errores introducidos a posteriori, como en el caso del Pro 35, otros adaptadores u ópticas en general.
- ❖ **El nivel de negros:** Establece el punto donde comienza a construirse la señal de vídeo. Crea un negro perfecto de referencia (ni agrisado, ni ultra negro).

- ❖ **Balance de Negros y Blancos:** Asegura que no hayan dominantes indeseables en grises, blancos y negros. En el caso de los blancos, a los píxeles que la cámara identifica como blancos (luz=1) los corrige de forma que no estén "teñidos" de color. Para ello compara el Canal Verde (G) con los canales (R y B) e iguala éstos hasta que los tres tienen la misma amplitud. Con el balance de negros pasa lo mismo. Electrónicamente estas cámaras trabajan a 3200°K, lo que supone otro punto sin retorno sobre la imagen original RGB.
- ❖ **El white Clip/ Knee /Slope:** Extiende el rango dinámico en la zona de las altas luces, preservando más detalle.
- ❖ **La ganancia electrónica:** Permite forzar la matriz de luminancia para ganar mayor sensibilidad... Y ruido electrónico, por supuesto.
- ❖ **La matriz:** Amplía el rango de matices de color que la cámara capta.
- ❖ **El detalle:** Mejora la sensación de nitidez en la imagen, mediante un ultra negro-ultra blanco en los bordes de la imagen. Actualmente no tiene mucho sentido en HD, ya que tanto el sistema óptico como la captura garantizan una más que adecuada calidad de imagen.

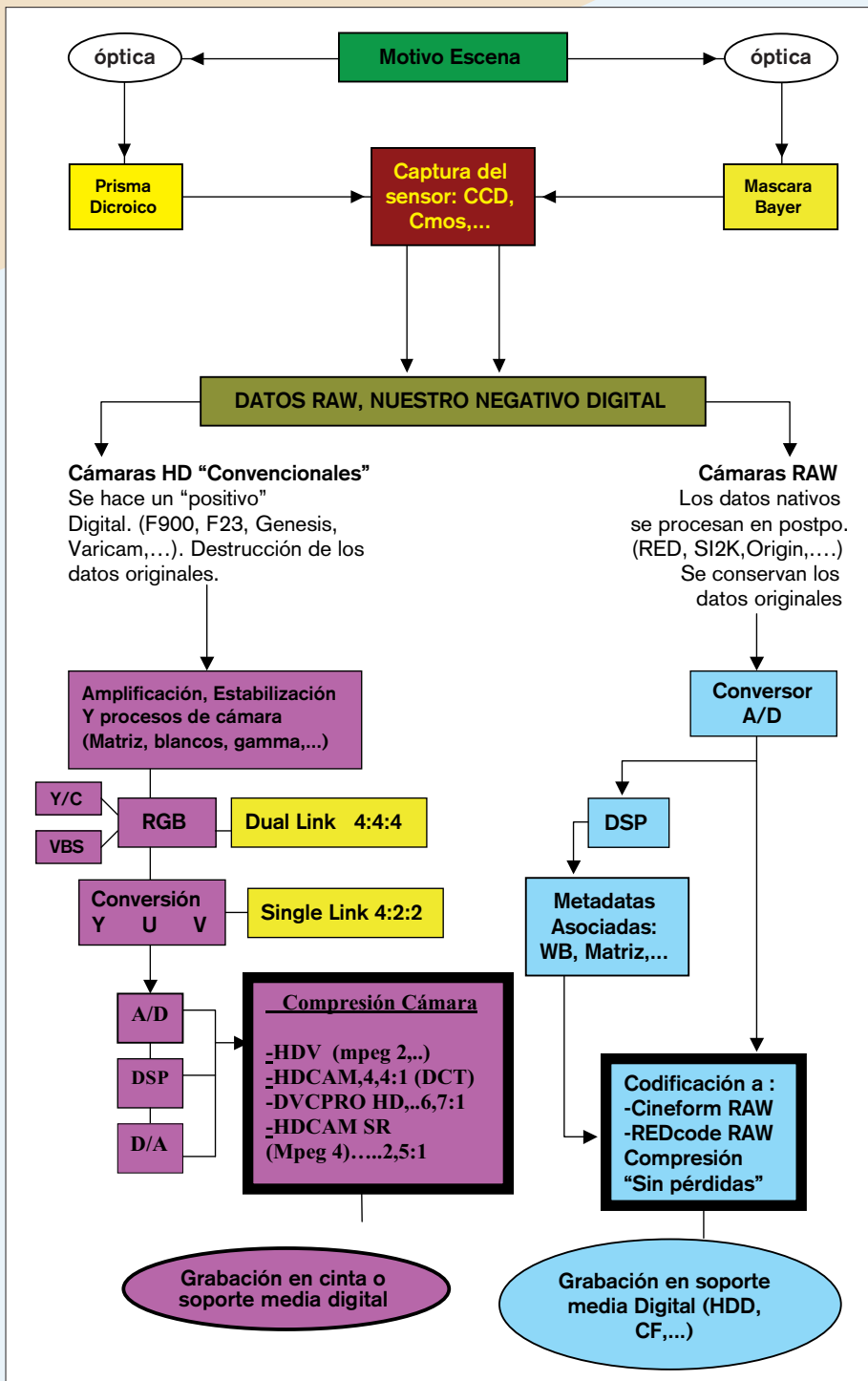
Todos estos parámetros vienen establecidos en preset de fábrica. El usuario puede acceder a ellos mediante los menús de cámara. Comentar que en los sistemas donde sí hay compresión (HDCAM, DVCPRO HD...) es posible llegar a medio camino en cámara sin dejarlo todo para posproducción, eso sí, sabiendo las limitaciones de forma que lo que se toca sea recuperable. Esto es lo que hacen a diario operadores de CCU en diferentes espacios televisivos.

Nota: Desactivar items como matriz o detail no quiere decir que anulemos estos circuitos de cámara, lo que hacemos es no darle a la cámara una referencia estándar basada en una norma de color (ITU-709, SMPTE, NTSC...)

Una vez que las tres señales RGB están suficientemente amplificadas y corregidas, es el momento de convertirlas al sistema de vídeo por componentes, máximo exponente (después del RGB) en cuanto a calidad de vídeo nos referimos. Esta señal acabará siendo digitalizada (HD-SDI) o se codificara a PAL, NTSC en el caso de la definición estándar. Reseñar que aún hay cámaras de vídeo que usan sistemas de vídeo compuesto, y de vídeo separado (Y/C).

Sistema de video por componentes

Las señales R, G y B son introducidas en la matriz de luminancia y sumadas en la siguiente proporción:



$Y=0,29(R) + 0,58(G)+0,11(B)$. Así obtenemos "Y" (luminancia).

Posteriormente, se invierte la luminancia, obteniendo la señal "-Y", que a su vez será sumada con las señales R, B en la matriz de crominancia.

Como curiosidad, tanto la señal de vídeo por separado como la de componentes van separadas en dos pistas, una para la luminancia y otra para las señales R-Y y B-Y, pero en el caso de la señal por componentes, no están mezcladas en la misma pista, sino que una va a continuación de la otra.

En la señal vídeo compuesto, la de menor calidad de las tres, toda la información de color y luminancia van juntas en una misma pista.

3.2 Proceso típico cámara RAW

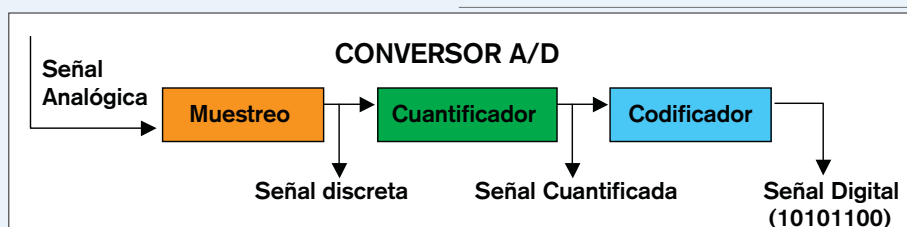
El concepto Raw (crudo) se basa en un formato de archivos digitales compuesto por todos los datos que el sensor puede capturar a partir de la imagen que llega a través de la óptica. Son ficheros con tamaños muy grandes que, como los casos del redcode Raw (codec de la Red One) y el Cineform Raw (codec de la SI2k), pueden ser comprimidos para reducir su tamaño. De hecho, en un futuro próximo, desde Red aseguran que estará operativo el puerto Raw para poder capturar toda la información del sensor sin ningún tipo de compresión. Esta compresión mínima no es comparable a la que sufren las señales de vídeo digitales en componentes 4:2:2, incluso el RGB 4:4:4, puesto que hablamos de datos informáticos puros.

Los datos Raw captados en bruto por el sensor constituyen un auténtico "negativo digital", y lo más importante: no sufren una degradación o manipulación por la cámara. Posteriormente, a través de un software podremos abrir estos archivos y procesarlos correctamente.

El diseño del sensor (resolución, profundidad de color...) junto con procesos posteriores definirán la calidad final de la cámara. Programas como el Red Alert, Redcine o el DVR Silicon permiten recuperar el Raw, sustraer la información de color de los datos nativos o hacer correcciones de color a posteriori, entre otras. Para ello aplican un "debayering", posterior a la "descompresión" o durante la captura.

4. El Conversor A/D

El conversor A/D transforma la señal origen analógica (continua en tiempo y amplitud) en una señal digital (discreta) en la cual, a partir de la información eléctrica, a cada píxel se le asigna un determinado valor fijo formado por unos y ceros (sistema binario). Estos valores serán interpretados como un color diferente en



los canales RGB que forman la imagen final.

Las ventajas de hacer esto son:

- ❖ La señal digital puede ser reconstruida con más facilidad que la analógica, recuperando y regenerándose sin introducir distorsiones típicas de la señal analógica.
- ❖ Una mayor facilidad de procesamiento de la señal en software de edición, corrección de color, efectos...
- ❖ Multigeneración infinita de copias digitales, sin pérdida de calidad, en soportes media tales como discos duros, tarjetas de memoria sólida, discos ópticos... En el caso de la cinta magnética (digital) sí que hay pérdida, aunque mucho menor que la cinta magnética analógica (soporta 4 ó 5 generaciones)
- ❖ El proceso básico de esta conversión es el coger muestras de la amplitud de la señal (analógica) de forma periódica y traducirlas numéricamente al sistema binario. Durante esta conversión intervienen cuatro procesos: Muestreo, Retención, Cuantificación y codificación.

Nota: En el caso de la captura en Raw no podemos hablar de una señal concreta, como puede ser la de componentes analógicos de una cámara de vídeo convencional. En este caso, el proceso A/D se hace en base a los datos en crudo del sensor.

4.1 Muestreo

En esta primera etapa se adquieren muestras de la amplitud de onda de forma periódica (recordemos que la señal aún es analógica). El número de muestras recogidas por segundo es lo que se conoce con el nombre de "frecuencia de muestreo". Es aquí donde podemos distinguir de nuevo dos caminos diferentes entre la captura en Raw y el HD convencional. En Raw se recogen todas las frecuencias origen de la señal mientras que en HD no. El muestreo típico 4:2:2 del HD convencional (Betacam digital, Dvcpro HD,) hace referencia a que se toman 13.500.000 muestras por segundo (13,5 MHz) para la luminancia (Y) y la mitad, es decir 6,75MHz, para las otras dos componentes que forman la crominancia; R-Y / B-Y. Como resultado se expresa 4:2:2 (donde 4= "Y" todas las muestras, 2= R-Y la mitad de muestras, 2=B-Y la mitad de muestras). En el caso del 4:4:4 la cosa mejora sustancialmente, ya que teóricamente se

recogen el mismo número de muestras para la crominancia y luminancia, no obstante y comparándolo con el Raw, éste sigue siendo mejor ya que ni siquiera es RGB 4:4:4 (como la F23/F35 o la Genesis) pues conserva todos los datos en bruto a la salida del sensor.

4.2 Retención

Una vez se han tomado las muestras, éstas son retenidas para que el siguiente proceso sea capaz de evaluarlas correctamente (cuantificadas).

4.3 Cuantización

En este siguiente proceso se mide el nivel de voltaje de cada muestra y se le asigna un valor (con mayor o menor margen) cuyo resultado sea único al nivel de salida. Por lo que existe una diferencia entre la señal eléctrica analógica y (desde este momento) la señal digital que acabamos de cuantificar. Esta diferencia al ser reproducida previo paso del decodificador se convertirá en ruido.

Para reducirlo se pasa la señal por un compresor logarítmico antes de su cuantización, ya que así hay menos variaciones en la muestra de voltaje resultante. El "dither" sacrifica relación S/R para impedir que al ser cuantificada (la señal) ésta muestre patrones aleatorios de ruido. Esta reducción es de aproximadamente 6 db. Este proceso es sumamente importante cuando se convierten señales ya digitalizadas (como de 10 a 8 bits / en la Varicam, F900, 750...) en procesadores digitales de señal (DSP).

4.3.1 Rango dinámico

Aparte de los característicos ruidos generados por los diversos tipos de sensores, también se generan estructuras de ruido estático, derivado de los procesos de muestreo y retención.

Este ruido se conoce como "ruido de cuantificación" y forma parte de la base inferior de "ruidos" que nos limitan el rango dinámico en el diseño del sensor. Este no es ni uniforme ni fijo y es debido a que la corriente de oscuridad varía de un píxel a otro, es decir, no es lineal, por lo que cada uno de los píxeles que integran un sensor tiene su propio y único patrón fijo de corriente de oscuridad. Dependiendo de cómo se fabrique y posteriormente eualice, se conseguirá minimizar más o menos. Normalmente se

añade un BIT (aprox. 6 db) de seguridad, y lo que se hace, por decirlo de alguna manera, es igualar todos los ruidos inherentes al diseño del sensor y su posterior muestreo de voltaje en cada píxel.

Este tipo de "ruido" es la base para determinar el rango dinámico (latitud) que tendrá el sensor en la parte baja de la señal (negros y sombras). Para la parte superior (altas luces, tonos medios), el límite lo determina la capacidad de los píxeles para acumular cargas. Si tomamos como referencia un nivel "correcto" de exposición, esta capacidad sería el número de niveles de luz capaces de ser manipulados por los elementos activos del sensor. Si vamos doblando la apertura del diafragma en relación a un gris medio, por ejemplo, llegará un punto (1, 2, 3, 4 stops) donde el sensor no es capaz de "acumular carga útil". Por tanto, podemos decir que la "señal útil" existente entre los diferentes tipos de ruido (**corriente de oscuridad, ruido térmico, ruido de cuantización**) y la capacidad del píxel para acumular cargas nos dan el rango dinámico de la cámara.

Esta relación señal-ruido (S/N) viene expresada en db. A un mayor número, mayor calidad. Es posible asociar que cada 6 db de S/N hay 1 BIT de profundidad de color, y que cada BIT nos da 1 stop de rango dinámico (latitud). Esta asociación es posible, ya que los sensores, a diferencia del negativo, trabajan de forma lineal, por lo que esta relación se considera correcta. En un supuesto sistema de 12 bits, tendremos 72 db de rango dinámico, lo cual nos dará, de forma orientativa, 12 stops de latitud. Hay que tener en cuenta el hecho de añadir 1 BIT de seguridad (6db) por lo ya comentado en relación al ruido de cuantización.

Una vez que tengamos el ruido de cuantización más bajo que el ruido de la señal analógica muestreada, será suficiente para enviarla al siguiente proceso (la codificación).

4.4 Codificación

En el codificador, los valores obtenidos de la cuantización pasan a valores de código binario. En función de un sistema u otro se aplicará un método de codificación específico, conocido como Codec (abreviatura de codificador-descodificador). Cada codec es particular de uno u otro sistema, y aunque guardan similitudes, a la hora de definirlos se podrían establecer características o parámetros a tener en cuenta comunes a todos ellos:

- ❖ **Frecuencia de muestreo:** Son las cantidades de muestras tomadas en un instante de tiempo (en el proceso de muestreo) que servirán para reconstruir la señal comprimida.

- ❖ **Resolución** (profundidad de color): Recordemos que cada elemento activo de la imagen, es decir cada píxel, tiene el poder de transformar el punto de luz origen en una reacción eléctrica que nos dará mayor o menor información de valores de grises en función de la capacidad que tengan en poder discriminar más o menos información. Una vez que la señal ha pasado por el codificador, esta capacidad analógica de discriminar información por parte del píxel es transformada digitalmente a un valor binario. A un mayor número de bits, una mayor capacidad de convertir información original.

El número de bits nos dará la mayor o menor capacidad de reproducción de la señal original. Cuantos más bits (8, 10, 12...) más precisión por píxel / muestra. Sirva como ejemplo un sistema de 12 bits, en el cual se obtendrían 4096 valores entre el negro y el blanco. Para entenderlo mejor, en un supuesto sistema de 8bits, cada muestra puede ser representada por una combinación de 8 dígitos. 00000000= equivaldría al blanco, 11111111= equivaldría al negro. Como 8 elevado a la 2 son 256, obtendríamos 256 valores posibles de grises o colores (00000001, 00000011...). Siguiendo en el supuesto de los 8 bits, y teniendo en cuenta que hay tres canales RGB en vídeo, obtendremos 24 bits, es decir 16 millones de grises/colores (256x256x256) con lo que las posibilidades de reproducir la escena original suben considerablemente y, con ellas, la calidad de la imagen captada, muy parecida a la del ojo humano. No obstante, los 8 bits se quedan cortos, sobre todo cuando se trata de captar imágenes en movimiento que posteriormente han de ser corregidas de color o editadas, por lo que actualmente la tendencia es implementar la grabación a 10 y 12 bits (lineal o log), más adecuada a la búsqueda del máximo exponente de calidad.

Este número de muestras representadas (8, 10, 12 bits) determinan el rango dinámico resultante de la S/R, el cual, a su vez, viene limitado, como vimos antes por el ruido de cuantización. A un mayor número de bits, mayor rango dinámico y, por supuesto, mayor S/R en cuantización.

- ❖ **BIT Rate:** El BIT rate es la velocidad o tasa de transferencia de datos. Su unidad es el bit por segundo (bps). Esta velocidad puede ser constante o variable. En la constante se cuantifica toda la señal por igual, mientras que en la variable (Red, SI 2K...) se tienen en cuenta zonas de la imagen con más o

menos información, siendo ésta mucho más eficaz. Si imaginamos un coche en fondo croma, en un BIT rate constante se daría la misma importancia a cada píxel del croma, aumentando enormemente y de forma innecesaria la memoria para almacenar ese frame. Si usamos un BIT rate variable, todo el croma será agrupado como una unidad reduciendo mucho la memoria necesaria para cuantificar ese "frame".

- ❖ **Pérdida:** En la mayoría de codecs, al hacerse la compresión se pierde información, es decir, que es imposible reconstruir la señal original a partir de la comprimida. En estos casos hablamos de compresión con pérdidas (mpeg2, mpeg4...).

Mediante algoritmos complejos de codificación es posible reconstruir la señal original (sin pérdidas / lossless). Lo que se hace es agrupar, como vimos antes, toda la información eliminando la repetida y agrupándola para que ocupe menos. En el caso de las cámaras de vídeo SD, se suelen emplear codificadores PAL, NTSC, a partir de la señal por componentes analógicos, donde se usa el subsampling y cuantificación antes de la codificación para reducir el tamaño de la señal. Por otro lado, en el caso del HDCAM, de las 1920 x1080 líneas originales, pasamos en compresión a 1440 x 1080 líneas. En el TV final las imágenes decodificadas son interpoladas en la pantalla y se perciben casi idénticas. La vista humana es más sensible a cambios de luminancia que de crominancia, por eso el HDCAM (MPEG) muestrea dos de crominancia por cuatro de luminancia.

5. DSP

Antes de que la señal sea comprimida y finalmente grabada, pasará por una serie de microprocesadores, compuestos básicamente de un hardware y software, los cuales realizan los cálculos necesarios para procesar (nuevamente mediante algoritmos) señales de vídeo en tiempo real. Los procesadores DSP son habituales dentro de una cámara de altas prestaciones y en muchas ocasiones se interponen entre el conversor A/D y el posterior conversor digital analógico.

Los DSP necesitan una memoria para almacenar los datos y un programa específico de cada fabricante para ejecutar la aplicación. Son procesadores complejos que tienen la capacidad de trabajar con varios datos en paralelo. Una de sus funciones típicas es la de poder manipular la señal en tiempo real (CCU, remotos...) y

tocar así los diversos procesos de la señal de cámara (gamma, knee...).

La Red One también cuenta con procesadores DSP, cuya función principal es la de corregir ya en cámara (sin afectar al RAW) las previsualizaciones de las diferentes salidas de cámara. Es posible variar el balance de blancos, o la distribución de la ganancia de cámara, siendo estos valores almacenados como metadatos dentro de los clips creados por la cámara. Posteriormente, o bien los recuperaremos o bien los descartaremos para postproducción.

El caso de la SI es similar, ya que incorpora un software dentro de la propia cámara que hace esta función, mediante algoritmos aplicados sobre el debayering del sensor, siendo siempre recuperables los datos capturados en Raw.

Otras funciones de los DSP pasan por hacer una corrección de gamma, restaurar el color desde la máscara Bayer, preparar la compresión o simplemente ajustar niveles automáticos de blancos, ganancia, contraste, en cámaras de vigilancia.

6. Compresión

Una vez tenemos la información codificada (datos informáticos, señal digital,) lo que necesitamos es reducir su volumen o ancho de banda, antes de grabar o emitir, de forma que se grabe toda (o casi toda) empleando el mínimo espacio posible. A este proceso, normalmente hecho en tiempo real, le llamamos compresión. Recordemos que la información que hemos codificado es la resultante de la cantidad de muestras y su posterior conversión como profundidad de color (resolución en bits). Como ya hemos visto antes, mediante algoritmos matemáticos intentaremos reconstruir la señal original. La finalidad de estos algoritmos es la de reducir en la medida de lo posible tanto la redundancia espacial como la temporal, de forma que al decodificar la imagen se minimice al máximo la distorsión visual y de paso consigamos un flujo de datos (BIT rate) cuanto menor, mejor.

El método más interesante es la compresión sin pérdidas, más acorde a la tecnología actual. En ella se transmite toda la información, pero eliminando la información repetida, agrupándola para que ocupe menos.

Cámaras como la Red o la SI 2K están basadas en wavelets, lo que permite procesar la imagen completa sin problemas de distorsión. Los wavelets "descomponen" señales en distintas bandas de frecuencias, lo

que es más acorde a las imágenes naturales y a la visión humana, ya que en éstas hay cambios bruscos en zonas homogéneas (como una transición sol-sombra). Los wavelets son un conjunto de procedimientos matemáticos que significan "pequeñas ondulaciones" y nos permiten comprimir una mayor cantidad de datos que se utilizarán posteriormente para almacenar la imagen. Así, será posible almacenar una imagen más detallada en un espacio menor. La Red o la SI 2K utilizan este tipo de compresión sin pérdidas basadas en wavelets. Siguiendo con el ejemplo de la SI 2K, lo que la cámara hace es codificar/comprimir a 10 bits Raw aunque de saque, el conversor A/D nos dé 12 bits.

La compresión se anota en términos de relación 2:1, 5:1. Una buena compresión es la relación de 2:1, de cada 100 megas (por cada 25fps) lo bajamos a 50.

Mediante el uso de codecs se reduce cada uno de los fps, según el sistema, a relaciones de 2:1 en el caso de la Sony F950; 4,4:1 para la Sony F750 o de 6,7:1 para la Varicam de Panasonic.

La resolución, frame rate, muestreo o codificación, serán algunos de los factores determinantes con los que, mediante operaciones matemáticas, llegaremos a la cifra final que nos indicará la frecuencia de datos (el bitrate) expresada en megabites. Esta es la cantidad de información que encierra el sistema, el flujo de datos que mueve la cámara.

La frecuencia de datos del HD es de 100-140 mbs, en comparación con la de los sistemas HDV que son de 25 mbs. A un mayor número de mbs una mayor calidad de imagen.

7. Debayering y demosaicing

La imagen captada por el sensor son escalas de grises, monocromas, siendo posteriormente, mediante el demosaicing y la aplicación de un W/B y matriz de color, cuando podremos visualizar correctamente la imagen. Las cámaras que incorporan un único sensor (CMOS o CCD) presentan tres filtros en forma de mosaico, formado cada uno por los colores primarios RGB (filtro Bayer). Sobre el sensor se superponen estos tres filtros de forma que cada uno de ellos deja pasar sólo su propia frecuencia (longitud de onda). El Rojo no tendrá frecuencias verde ni azul. Posteriormente, se reconstruirá la señal completa (RGB) a partir de los

píxeles adyacentes de cada color. Como curiosidad el filtro verde tiene el doble de microfiltros que los otros dos (R y B), ya que así el sensor "imita" la respuesta más sensible al verde del ojo humano.

En este proceso se generan diferentes matrices que toman como referencia los píxeles adyacentes y ayudan a discriminar cada matriz para reconstruir la información de color. De esta forma se consigue que cada píxel tenga la información con tres valores de color cada uno el conversor RAW rellena los espacios en base a la información de estas matrices. El debayering introduce ruido, aunque se va mejorando, muestra de ello es la Red, que al mejorar el algoritmo del debayering se logra reducir.

Durante estos procesos es normal la replicación de píxeles, interpolación bilineal y la mediana de interpolación. Otra de las cosas que hace el proceso de demosaicing es la de "recuperar" toda la resolución espacial del sensor y mostrarla en un espacio de color RGB, donde se puede ver la imagen más que aceptable.

Aunque el debayering sea inherente al software de cada fabricante y formato RAW, la extracción del color es un proceso de cálculo que también se puede realizar en el interior de la cámara mediante circuitos DSP, permitiéndonos realizar la operación en tiempo real, obteniendo de estos archivos una señal analógica o digital en RGB. A estos archivos vienen asociados una serie de metadatos que posibilitan poder abrir el Raw en un espacio de color tal como el de un monitor de rodaje, RGB con curvas de gamma de video (Itu-709). Estos metadatos nos permiten previsualizar en rodaje la señal más o menos como quedará.

Cámaras como la SI 2K, junto con el software de la cámara, permiten aplicar un demosaico sobre la marcha, en tiempo real directo (con los DSP) de la descompresión de sus archivos Cineform RAW. En este caso puedes aplicar la mitad de resolución V y H (quadlet) o la total (bilineal). La Red One hace un debayering (quadlet) en previsualización mediante procesadores DSP que nos permite ver cómo quedará más o menos la imagen sin necesidad de aplicar software en rodaje (Red Alert), aunque también está la opción de hacerlo en el mismo set de una manera más completa con este tipo de software.



Gabi García es un foquista especializado en cinematografía digital con amplia experiencia tanto en publicidad como ficción. Fue uno de los primeros ayudantes de cámara en trabajar en HD previa migración del SD. Ha dado clases en las principales escuelas de cine y es colaborador activo del TEA desde su primer año.