

5

Codificación digital de la imagen

Francisco UTRAY

*Profesor de Comunicación Audiovisual de la Universidad Carlos III de Madrid.
Grupo de investigación TECMERIN.*

Luis OCHOA

Colorista y diseñador de flujos de trabajo y director de Red 709 y 709 MediaRoom.

Sumario:

5.1	La imagen digital	122
5.1.1.	<i>El archivo informático de imagen</i>	122
5.1.2.	<i>Resolución espacial</i>	123
5.1.3.	<i>Profundidad de color</i>	123
5.1.4.	<i>El canal alfa</i>	125
5.1.5.	<i>Tamaño del archivo en Bytes de una imagen fija</i>	126
5.2.	Resolución espacial en cine digital y televisión	127
5.2.1.	<i>Los formatos de vídeo para televisión</i>	127
5.2.2.	<i>Resolución espacial en cine digital</i>	129
5.2.3.	<i>Convergencia de los formatos de cine digital y la televisión</i> 129	
5.2.4.	<i>Proporción del píxel</i>	130
5.3.	Resolución temporal	131
5.3.1.	<i>Frecuencia de presentación de los fotogramas</i>	132
5.3.2.	<i>Exploración entrelazada y progresiva</i>	132
5.3.3.	<i>Masterización internacional y resolución temporal</i>	133
5.3.4.	<i>Variación de la resolución temporal y de la velocidad</i> ...	135
5.4.	Compresión de la señal de vídeo	136
5.4.1.	<i>Flujo de transferencia binaria</i>	136
5.4.2.	<i>Señal de vídeo RGB y YUV</i>	137
5.4.3.	<i>Submuestreo de color</i>	138
5.4.4.	<i>Compresión espacial y compresión temporal</i>	139

5.5. Formatos de archivo y códecs.....	140
5.5.1. Códecs, contenedores y formatos de vídeo	141
5.5.2. Comparación de dos formatos de vídeo	144
Referencias bibliográficas.....	146
Lecturas recomendadas.....	146

5.1 La imagen digital

5.1.1. El archivo informático de imagen

Un archivo o fichero informático, es un conjunto de bits almacenados en un dispositivo. El archivo informático de una imagen contiene la codificación binaria (en bits) de la información de brillo y el color que componen dicha imagen.

La unidad mínima de la imagen digital es un píxel. Este término, ya castellanizado, proviene de la contracción de *picture element*. Cuantos más píxeles compongan una imagen, mayor posibilidad de representar los detalles. El número de píxeles que componen una imagen se denomina ‘resolución espacial’ y es un indicador técnico de la calidad fotográfica.

La resolución espacial es uno de los factores que afectan al proceso de codificación digital ya que el archivo informático de imagen tiene que registrar la información sobre brillo y color de cada uno de los píxeles de esa imagen. Por lo tanto, cuanto mayor sea la resolución espacial de la imagen, mayor será también el tamaño de archivo, es decir, el ‘peso’. Expresado con otros términos, el número de bits necesarios para la codificación binaria.

Pero la cantidad de píxeles que forman la imagen digital no es el único factor que determina el peso del archivo y la calidad técnica de la imagen. También se tiene que considerar la cantidad de bits que se utilizan para describir el brillo y el color de cada uno de esos píxeles. Es lo que se denomina ‘cuantificación’ o en términos fotográficos, profundidad de color.

El peso, es decir el tamaño del archivo informático de imagen, que se mide en Bytes (B), kilobytes (KB), megabytes (MB) o gigabytes (GB). Un Byte (B) es una unidad de medida informática que corresponde a 8 bits. Los prefijos kilo, mega y giga corresponden respectivamente a miles, millones, y miles de millones.

Otro factor determinante para el tamaño en Bytes de una imagen es el sistema de compresión que se utiliza en la codificación. Es decir los algoritmos que se han utilizado para optimizar su tamaño y el espacio necesario para su almacenamiento o para la transmisión.

En este capítulo veremos con cierto detenimiento los conceptos técnicos que afectan a la codificación digital de las imágenes y que son válidos tanto para el ámbito de la fotografía como para el vídeo o el cine digital.

5.1.2. Resolución espacial

La resolución espacial indica las dimensiones en píxeles de una imagen, es decir, el número de píxeles que constituyen su altura y anchura.

En una cámara fotográfica digital, la resolución espacial indica la cantidad de información escaneada por el sensor y determina la capacidad de la imagen para reproducir el detalle y las texturas del objeto fotografiado. Este valor se puede expresar como el número total de píxeles que forman la imagen o mediante dos cifras: el número de píxeles en sentido horizontal y vertical. Así, cuanto nos referimos a una imagen de 1.920 píxeles horizontales por 1.080 verticales estamos ante una imagen de 2.073.600 píxeles de resolución espacial. Este dato también se expresa como 'una resolución de 2 megapíxeles', es decir, aproximadamente dos millones de píxeles.

Pero en entornos profesionales de producción de vídeo y postproducción se emplea generalmente la forma más precisa 1920x1080.

En cine digital es habitual referirse a la resolución espacial indicando únicamente el valor aproximado de píxeles en sentido horizontal. Los valores normalizados de resolución espacial para la exhibición de cine digital son 2K y 4K, indicando así que la imagen está compuesta por un valor aproximado de 2.000 o 4.000 columnas de píxeles. El apartado 5.2 está dedicado a la resolución espacial en cine y televisión.

En términos de codificación binaria, la resolución espacial indica el número de muestras (*samples*) que hay que codificar para generar un archivo digital. De ahí surge el término 'frecuencia de muestreo' (*sampling rate*) que indica el número de muestras (píxeles) que se toman por segundo en un flujo de imágenes en movimiento.

5.1.3. Profundidad de color

Otro indicador fundamental de la calidad de una imagen digital es la profundidad de color (*color depth* o *bit depth*). Con este dato se expresa cuánta información del color está disponible para mostrar en pantalla o imprimir cada píxel de una imagen.

Cada píxel puede ser codificado con un número variable de bits. El bit es la unidad mínima de la codificación binaria que solo puede tener dos estados: uno o cero; encendido o apagado. Por lo tanto si solo disponemos de un bit para describir un píxel podremos usarlo para indicar si es blanco (1) o negro (0). Con dos bits se pueden generar cuatro combinaciones (00, 01, 11 y 10) que se pueden utilizar para designar el negro, el gris oscuro, el gris claro y el blanco. Si disponemos de más bits para describir el brillo de un píxel podremos utilizarlos para expresar toda una gama de niveles de gris.

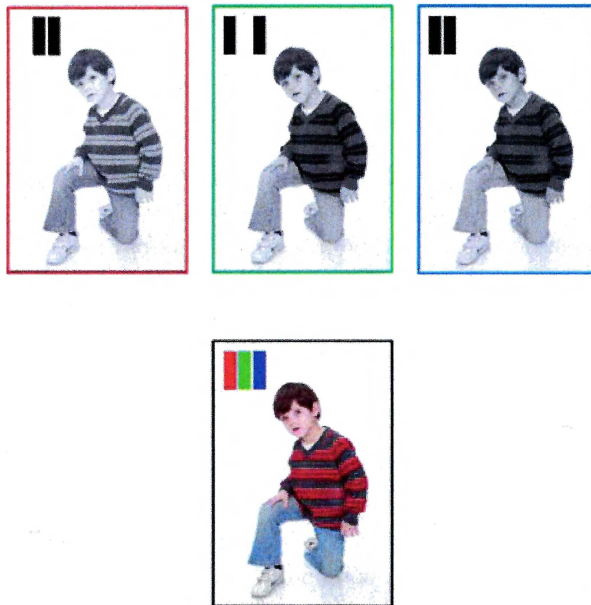
Las imágenes en color se pueden descomponer en los valores de brillo de cada una de las componentes de los colores primarios rojo, verde y azul (RGB, del inglés *Red*,

Blue, Green). La mezcla de estos tres colores, en distintas proporciones, puede generar todo el espectro cromático visible. La utilización de varios bits para cada componente de color RGB, nos permitirá alcanzar la precisión de color necesaria que requiere una imagen de calidad fotográfica.

El número de bits que se utiliza para cada componente de color se conoce como la profundidad de color o la profundidad de bits (*component bit depth*). Mayor profundidad de color (más bits de información por píxel) significa más matices de color disponibles y por lo tanto una representación del color más precisa. Es el valor de la 'cuantificación' del proceso de codificación digital.

En la siguiente figura se muestra una imagen en color y su descomposición en los tres canales primarios. La mezcla de estas tres imágenes monocromáticas filtradas con su correspondiente color primario genera una imagen en color con todos sus matices cromáticos.

Figura 5.1. *Descomposición de una imagen en sus tres componentes de color RGB*



Fuente: *elaboración propia*

La profundidad de color más habitual es de 8 bits por cada canal RGB, es decir una imagen de 24 bits: 8 bits para la componente roja, 8 para el verde y 8 para el azul. Es lo que también se conoce en los entornos informáticos como color verdadero (*true color*) porque es la resolución de color que necesita una imagen para tener una apariencia

fotográfica estándar. La codificación con valores inferiores de profundidad de color (8 o 16 bits) solo se usarán para gráficos o para pequeñas imágenes que sin altos requisitos de calidad. Pero actualmente, ya sea en fotografía, en vídeo o en cine digital, se pueden utilizar 10, 12 o incluso 16 bits para cada canal RGB. Estos valores permiten aumentar la gama de colores hasta niveles iguales o superiores a los que puede percibir el ojo humano y superiores también a los sistemas de monitorización de los que disponemos en los ordenadores.

5.1.4. El canal alfa

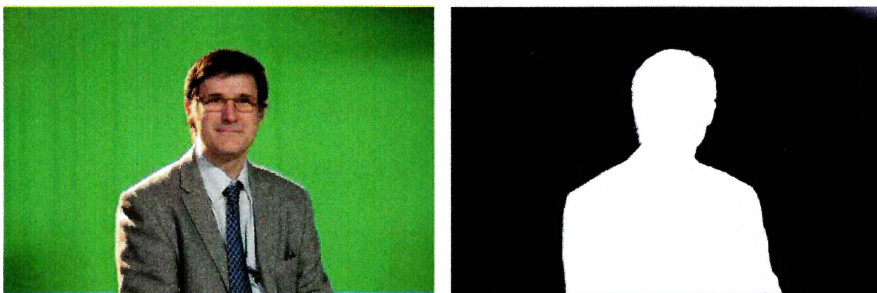
Algunas imágenes pueden contener un cuarto canal que se denomina 'canal alfa' (*Alpha Channel*). El canal alfa es una máscara que indica el nivel de transparencia de los píxeles. Blanco indica opacidad total, negro transparente y los distintos niveles de gris las semitransparencias.

El canal alfa se utiliza para la composición de imágenes por capas. Por ejemplo el icono de un logotipo que se va a 'pegar', o superponer, sobre una imagen, tendrá que llevar un canal alfa que delimita su forma.

Se representa mediante una imagen en blanco y negro y para la codificación también se pueden usar 8 bits. Por lo tanto las imágenes RGB + canal alfa codificadas con 8 bit por canal tendrán en total 32 bits de profundidad de color (8 bits por canal x 4 canales = 32 bits).

En la siguiente imagen podemos ver el canal alfa de una imagen que nos permitirá incrustar la figura (*foreground*) en cualquier otro fondo (*background*) sustituyendo la zona de color verde plano.

Figura 5.2. Representación gráfica del canal alfa



Fuente: elaboración propia, con una fotografía de Manuel Armenteros

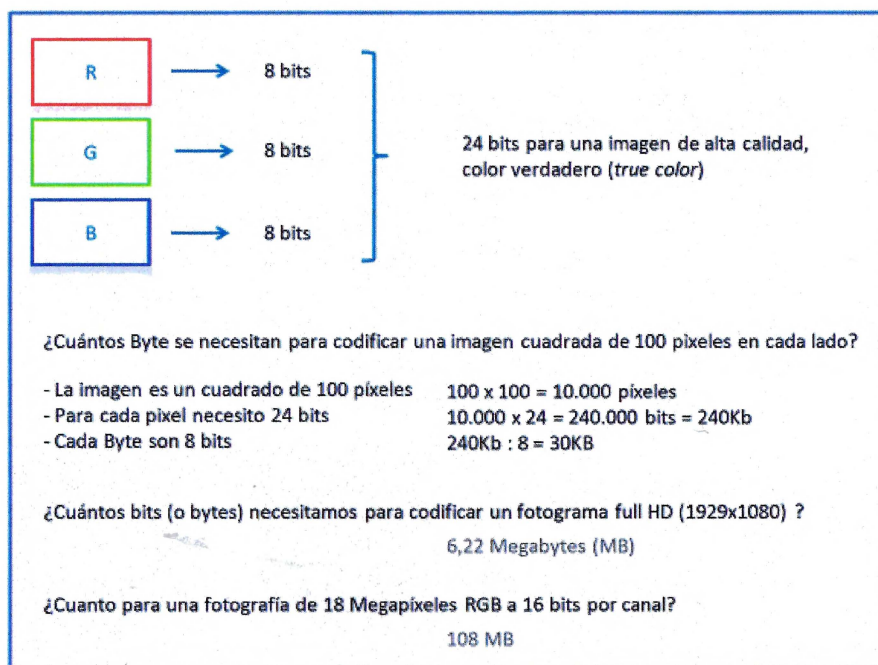
5.1.5. Tamaño del archivo en Bytes de una imagen fija

Para calcular el tamaño en bits que tendrá una imagen fotográfica sin comprimir de 8 bits por canal, bastará con aplicar una sencilla fórmula matemática. Tendremos que multiplicar la resolución espacial por el valor de la profundidad de color. En otras palabras, el número de píxeles que forman la imagen por el número de bits que empleamos para describir cada píxel. El resultado será el número de bits totales necesarios para la codificación.

El tamaño de las imágenes, sin embargo, no se suele expresar en bits, sino en Bytes. Un Byte está compuesto de 8 bits. Por lo tanto para calcular el peso del archivo en Bytes, tal y como nos lo muestra el ordenador, tendremos que dividir por 8 el resultado de la operación.

En la siguiente figura se propone un ejercicio de cálculo en el que será necesario aplicar esta fórmula matemática.

Figura 5.3. Cálculo de los Bytes necesarios para codificar una imagen fija sin compresión



Para terminar de comprender la codificación de la imagen falta aún considerar la compresión que veremos con mayor detenimiento más adelante.

A modo resumen, a continuación se listan los factores que determinan el tamaño en Bytes de una imagen:

- Número de muestras que componen la imagen: número de píxeles
- Tamaño de cada muestra: número de bits que se usan para definir el color de cada píxel en RGB. Es la profundidad de color expresada en bits.
- Formato y nivel de compresión. La compresión elimina información redundante o innecesaria en la codificación binaria de la imagen.

5.2. Resolución espacial en cine digital y televisión

5.2.1. Los formatos de vídeo para televisión

Para asegurar la compatibilidad de los equipos de captación de imagen, grabación, transmisión y recepción, desde los inicios de la televisión fue necesario establecer estándares tecnológicos internacionales. Mediante estas normas técnicas se han establecido los distintos modelos de televisión que se han utilizado en el mundo, concretando aspectos como la resolución de la imagen, la frecuencia de fotogramas por segundo o los procesos de exploración.

En los años 50 se desarrolló el primer estándar de televisión analógica en color, el NTSC, que se utilizó en Estados Unidos. En Europa se adoptó la norma PAL. A principio de los años 90 en Europa, el consorcio DVB (*Digital Video Broadcasting*) fue el encargado de definir el estándar para la radiodifusión digital. Pero igual que ocurrió en su día para las normas de televisión en color, los Estados Unidos crearon su propio sistema, el ATSC (*Advanced Television Systems Committee*). Japón también tiene un sistema propio, el ISDB (*Integrated Services Digital Broadcasting*). El resto de los países del mundo han adoptado uno u otro sistema según sus intereses y acuerdos comerciales.

En términos de resolución espacial de la imagen televisiva se identifican tres etapas históricas que significan un cambio sustancial de calidad: la resolución estándar (*SD standard definition*), la alta definición (*HD High definition*) y la tecnología aún en desarrollo de la ultra alta resolución (*Ultra High definition; Ultra HD; UHD*).

En la siguiente tabla se pueden ver los valores de resolución espacial y relación de aspecto de estos formatos siguiendo las normas europeas de televisión digital.

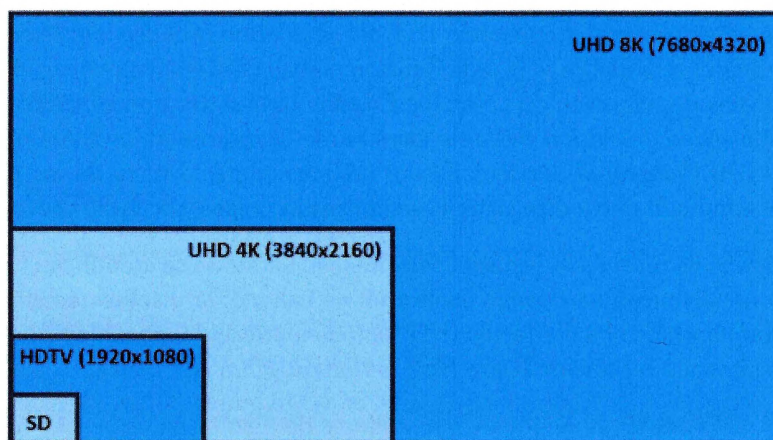
Tabla 5.4. Resolución espacial y relación de aspecto en las normas europeas

	Relación de aspecto	Resolución espacial
SD (<i>Standard Definition Television</i>)	4:3	720 x 576
HD (<i>High Definition Television</i>)	16:9	1280 x 720 1920 x 1080
4K UHD (<i>Ultra High Definition Television</i>)	16:9	3840x2160
8K UHD (<i>Ultra High Definition Television</i>)	16:9	7680x4320

Fuente: elaboración propia

En la siguiente imagen puede verse una representación a escala de la evolución en resolución espacial de la televisión, desde la resolución estándar (SD), a la alta definición (HD) y los formatos 4K y 8K (UHD).

Figura 5.5. Evolución de la resolución espacial de los formatos de TV



Fuente: elaboración propia

La televisión en sus inicios adoptó una relación de aspecto de 4:3 (1,33:1) siguiendo la tradición del formato académico del cine. Pero en los años 50 para la producción y exhibición cinematográfica se desarrollaron los formatos panorámicos consiguiendo así un elemento diferenciador y aumentando la espectacularidad del formato. En los años 2000, en televisión, aprovechando la migración digital, se implantó en todo el mundo el formato panorámico 16:9 (1.77:1).

Actualmente las grandes cadenas de televisión americanas y europeas emiten en los dos formatos de alta definición HD (1920x1080 y 1280x720). Pero aún es muy frecuente la radiodifusión en resolución estándar SD, tanto en su versión de proporción 4:3 (720x576) como en la de 16:9 (1050x576).

Los fabricantes de televisores están ya comercializando equipos de 4K Ultra HD, pero aún habrá que esperar para que este nuevo formato se generalice en la población mundial. Para los juegos olímpicos de Tokio 2020, los japoneses están haciendo pruebas con la tecnología de 8K.

5.2.2. Resolución espacial en cine digital

En los tiempos de la película fotoquímica, para la postproducción de cine con tecnología digital, era preciso escanear cada fotograma de la película en un archivo digital de alta calidad. Una vez procesado el material con sus efectos digitales, se generaba un nuevo negativo filmando de nuevo cada fotograma. El archivo digital resultante del escaneado se denomina '*digital intermediate*' (DI).

Los requisitos de calidad que se establecieron para el *digital intermediate* eran superiores a los del vídeo o la televisión de la época puesto que el material final se iba a difundir en pantalla grande en las salas de exhibición. El reto era aproximarse a la calidad de la película fotoquímica.

La empresa Kodak diseñó un sistema, denominado Cineon, que servía a este propósito. Se utilizó por primera vez en el año 1993 para la restauración digital del clásico de Disney *Blancanieves y los siete enanitos*. La película entera fue escaneada, fotograma a fotograma, en archivos digitales de 4K a 10 bits de profundidad de color y retocada con sistemas informáticos.

De esta forma se fueron estableciendo poco a poco los estándares técnicos para la migración digital de la industria cinematográfica. La resolución espacial se estableció en 2K o 4K, indicando de esta forma el número de píxeles aproximados en sentido horizontal (2.000 o 4.000). El número de píxeles en sentido vertical queda definido por la relación de aspecto (*aspect ratio*) del fotograma fílmico: ya sea 1,66:1, 1,85:1, 2,35:1, o cualquiera de los estándares utilizados por la industria del cine.

En la actualidad, los formatos normalizados para la exhibición en salas de cine digital tienen una resolución espacial de 2048x1080 para 2K y de 4096x2160 para el 4K y una relación de aspecto de 1,85:1 y 2,39:1.

5.2.3. Convergencia de los formatos de cine digital y la televisión

La evolución tecnológica del vídeo ha progresado atendiendo a los requisitos de la industria del cine para ofrecer una imagen de mayor calidad. En una primera etapa,

los cineastas empezaron a utilizar la tecnología digital solo para la postproducción. Posteriormente se desarrollaron las cámaras digitales y los proyectores para las salas. La digitalización de la televisión también abrió las puertas para la alta definición. En la actualidad, con los nuevos formatos de Ultra Alta Definición (UHD) se puede considerar que se ha consolidado la convergencia tecnológica entre cine digital y televisión.

En la siguiente figura se pueden seguir la cronología de esta evolución:

Figura 5.6. *Convergencia tecnológica del cine digital y la televisión de ultra alta definición*

	TELEVISIÓN	CINE
1990	SD, 4:3, 720x576	2K, 1,85:1, 1998x1080 2K, 2,39:1, 2048x858
2000	HD, 16:9, 1920x1080	4K, 1,85:1, 3996x2160 4K, 2,39:1, 4096x1714
2010	4K UHD, 16:9, 3840x2160	Formato emergente 8K 8K, 17:9, 8192 x 4320 8K, 16:9, 7680 x 4320
2020		

Fuente: *elaboración propia*

5.2.4. Proporción del píxel

Otro aspecto importante de cara a tener una visión completa de los formatos de vídeo digital es la proporción del píxel. Como norma general los píxeles son cuadrados (*square pixel, SP; full raster*). Sin embargo en algunos formatos de vídeo digital se utiliza un píxel rectangular (*non square pixel, NSP; thin raster*). Es lo que se conoce como proporción el píxel (*Pixel Aspect Ratio, PAR*).

El origen histórico de este fenómeno se encuentra en los acuerdos para la interoperabilidad de los formatos de vídeo norteamericanos y europeos en los inicios del proceso de digitalización de la televisión.

Los formatos que están afectados por esta característica son los siguientes:

Tabla 5.7. Formatos con píxeles rectangulares (thin raster)

			PAR
PAL	4:3	720x576	1,09
	16:9		1,45
NTSC	4:3	720x480	0,90
	16:9		1,21
HDV /DVCAM	16:9	1440x1080	1,33

Fuente: elaboración propia.

Todos los nuevos formatos tienen el píxel cuadrado. Sin embargo en producciones dirigidas a la televisión es aún muy frecuente trabajar con píxeles rectangulares.

En los sistemas de edición y postproducción se puede configurar sin dificultad este parámetro en las preferencias del proyecto.

Cuando hay que mezclar en un mismo proyecto planos con píxel cuadrado y píxel rectangular una opción es transcodificar una parte del material para que todo el proyecto tenga el mismo formato.

5.3. Resolución temporal

Los sistemas digitales que se utilizan hoy en día ofrecen cierta flexibilidad a la hora de definir la frecuencia de imágenes en la captación y en la reproducción. Cuantos más fotogramas tomemos de una acción más información tendremos sobre el movimiento que produce. Por ello la frecuencia de fotogramas por segundo (*fps*, *frames per second* o *frame rate*) de una secuencia se denomina 'resolución temporal' (*temporal resolution*).

La ilusión del movimiento de las imágenes animadas se produce por la combinación de dos fenómenos perceptivos: la fusión crítica del parpadeo y el movimiento aparente. La fusión crítica del parpadeo hace referencia a que nuestro sistema visual no percibe el parpadeo (*flicker*) cuando se produce a una frecuencia superior a 50 veces por segundo (50 Hercios). El segundo factor que genera la ilusión del movimiento en el cine es una característica de nuestro sistema visual, la percepción de movimiento, que también contribuye a generar esa sensación de continuidad (Bordwell y Thompson, 1995, pág. 3-4).

5.3.1. Frecuencia de presentación de los fotogramas

El movimiento en el cine se capta generalmente a 24 fotogramas por segundo (fps), alcanzando con esta frecuencia la sensación de continuidad. Sin embargo, para conseguir eliminar el parpadeo en proyección de película fotoquímica era necesario duplicar la frecuencia, presentando en la pantalla dos veces cada fotograma. De esta manera en proyección de cine se alcanzan 48 exposiciones por segundo eliminando así el efecto de parpadeo (*flicker*).

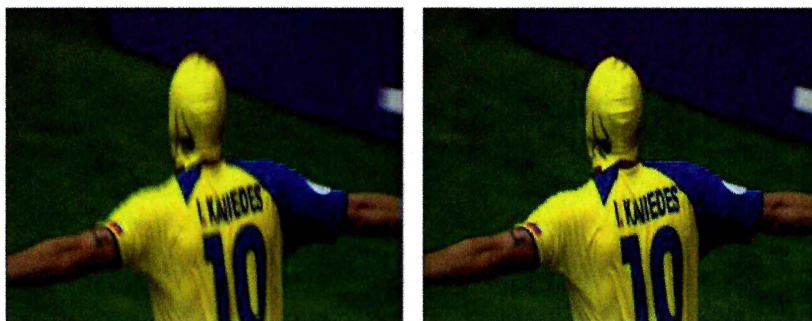
En televisión, la norma europea PAL optó por una frecuencia de 25 fps, y en el sistema americano NTSC utilizó 30 fps para la televisión en blanco y negro y 29,97 fps para la televisión en color. Mediante la tecnología del barrido entrelazado se consiguió duplicar esta cadencia alcanzando en el sistema europeo, 50 campos por segundo y el americano 60. En el siguiente apartado se desarrolla el funcionamiento de esta tecnología.

5.3.2. Exploración entrelazada y progresiva

En los inicios de la televisión, con la finalidad de aumentar la frecuencia de presentación de imágenes y evitar la percepción del efecto de parpadeo, las normas técnicas establecieron la división de cada fotograma o cuadro (*frame*) en dos campos (*field*), uno compuesto por las líneas impares de la imagen y otro por las líneas pares. Es lo que se ha denominado como exploración entrelazada (*interlaced scan*). Más adelante, con la aparición de los sistemas informáticos, se definió otra norma para la exploración de la imagen consistente en un barrido continuado, línea tras línea, de toda la superficie del fotograma o cuadro. Es lo que se denominó exploración progresiva (*progressive scan*).

En la siguiente figura, en la imagen de la izquierda, se puede apreciar el desplazamiento entre los dos campos en una imagen con mucho movimiento. La imagen de la derecha corresponde a uno solo de los campos y los bordes de la figura se aprecian nítidos.

Figura 5.8. Desplazamiento de los dos campos que forma el cuadro en la exploración entrelazada (imagen de la izquierda)



La exploración progresiva consiste en la exploración secuencial de cada línea de la imagen y el efecto parpadeo se compensa utilizando, en el equipo de visionado, ya sea un monitor profesional o el televisor doméstico, una frecuencia de 50 o 100 Hz de refresco de pantalla.

Tanto los monitores de ordenador como la mayoría de las pantallas de televisión de alta definición utilizan la exploración progresiva. Se puede afirmar por lo tanto que los entornos digitales han adoptado mayoritariamente el sistema de exploración progresiva. El aumento de dispositivos de visionado basados en exploración progresiva está permitiendo que todo el proceso de adquisición, procesamiento, transmisión y reproducción de las señales se haga con un mismo sistema de exploración, lo que está evitando efectos derivados de la conversión de entrelazado a progresivo.

Una desventaja de la exploración progresiva es que necesita, para la transmisión de la señal, mayor ancho de banda que la entrelazada. En otras palabras, el proceso de codificación digital de la exploración progresiva genera un flujo de bits superior. Este es el motivo por el cual las nuevas normas de televisión de alta definición siguen considerando válida la opción de la exploración entrelazada. Por ejemplo en Europa para la emisión de televisión en alta definición se utiliza el formato HDTV 1080i. La 'i' minúscula, viene de *interlaced*, e indica que se trata de un formato con exploración entrelazada. Cuando nos referimos a formatos con exploración progresiva, se utiliza una 'p' minúscula después del dato de resolución vertical (ejemplo 720p).

Las aplicaciones de edición y posproducción de vídeo permiten trabajar con ambas formas de exploración.

La exploración entrelazada hoy en día solo se utiliza en los entornos de radiodifusión televisiva, ya que tanto el cine digital como el vídeo en internet utilizan exploración progresiva.

5.3.3. Masterización internacional y resolución temporal

El mercado de los contenidos audiovisuales se ha desarrollado desde sus inicios en el contexto internacional y en el siglo XXI, más que nunca, las iniciativas de producción se conciben para ser comercializadas en todo el mundo y en todas las plataformas posibles.

El concepto de 'masterización internacional' implica la adaptación del montaje final –el máster– a los distintos estándares de resolución espacial, de relación de aspecto y de resolución temporal que existen en el mundo y en las distintas ventanas de explotación: cine, televisión e internet. La idea principal es trabajar el máster en el formato de mayor resolución para poder luego hacer los procesos de conversión a los formatos inferiores.

En relación con la resolución temporal hay que abordar la conversión a las distintas frecuencias de fotogramas (*frame rate*) que existen en el mercado internacional. Este proceso de conversión se denomina *pull down*.

Las frecuencias más habituales son: 25 fps, 29,97 fps, 24 fps, 23,98 fps y 48 fps. En la siguiente tabla se detallan los contextos en donde se usan:

Tabla 5.9. Valores más habituales de frecuencia temporal

25 fps	Es la frecuencia que se usa en la televisión PAL, en Europa y en las regiones del mundo donde la electricidad va a 50 Hz.
29,97 fps	Es la frecuencia que se utiliza en la televisión NTSC, en Estados Unidos y en las demás regiones del mundo donde la frecuencia eléctrica va a 60 Hz.
24 fps	Es la frecuencia que se utiliza tradicionalmente en cine.
23,98 fps	Es la frecuencia que se utiliza para el cine digital para facilitar las conversiones a la frecuencia 29,97 de la televisión NTSC.
48 fps	Es un estándar de resolución temporal opcional en la exhibición de cine digital.

Fuente: *Elaboración propia*

En algunos casos, cuando las diferencias son pequeñas, se puede hacer una conversión directa fotograma a fotograma. Pero se producirá una sutil aceleración o deceleración de la imagen y el consiguiente cambio de tonalidad en el audio¹ (*pitch*). Para evitar este cambio de tono, se puede modificar la velocidad de la imagen pero mantener el audio con su duración original. En este caso será necesario dedicar un importante esfuerzo en la resincronización de las pistas de sonido a lo largo de toda la película.

Pull down es la denominación que tiene el cambio de frecuencia (*frame rate*) cuando se aplican algoritmos para crear nuevos fotogramas por interpolación (*frame blending*). El *pull down* más habitual es cuando se pasa de 23,98 fps a 29,97 fps. Se utiliza para pasar el cine a la televisión en Estados Unidos.

Para la conversión de 24 fps a 25 fps, habitualmente se acelera la imagen y se trabaja manualmente la sincronización del audio, aprovechando para los ajustes los fragmentos de vídeo sin diálogos.

¹ Cuando se acelera la imagen y el sonido se produce en la pista de audio un cambio de tonalidad. Todos los sonidos se transportan a una frecuencia más aguda. Cuando se ralentiza ocurre lo contrario, se reduce la velocidad y el sonido se traslada a frecuencias más graves. En inglés se utiliza el término *Pitch* para designar estas modificaciones de altura musical del sonido.

Para una producción que se vaya a explotar en el mercado global lo más habitual es captar la imagen a 23,98 fps, después hacer el *pull down* a 29,97 fps para la televisión en Estados Unidos y convertir para Europa a 25 fps desde el 23,98 fps.

5.3.4. Variación de la resolución temporal y de la velocidad

Peter Jackson en la grabación de la película *El hobbit: un viaje inesperado* (2012) experimentó con una resolución temporal de 48 fps (*high frame rate, HFR*) en toda la cadena de producción y exhibición de cine. Tanto la captación de imágenes en cámara como la creación de gráficos 3D se produjo a 48 fps. La postproducción, específica para imagen estereoscópica (exhibición 3D) también se adaptó a esta frecuencia así como la proyección en salas. El objetivo era conseguir mayor calidad y fluidez en la representación del movimiento.

En las retransmisiones deportivas también es frecuente aumentar la frecuencia de captación para registrar con mayor detalle la acción de los deportistas. Es lo que se denomina 'cámara superlenta' (*super slow motion*) que, en realidad, es una cámara que trabaja a alta velocidad. Se utiliza en las repeticiones para reproducir el movimiento ralentizado con alta calidad. Existen en el mercado profesional cámaras que son capaces de grabar a velocidades muy altas (¡hasta miles de fotogramas por segundo!) para conseguir por ejemplo reproducir a cámara lenta una explosión o el vuelo de una bala de pistola.

Incluso las cámaras digitales DSLR que se distribuyen para el mercado doméstico permiten grabar vídeo en alta definición a velocidades superiores a lo normal (por ejemplo a 60 fps) para mejorar la representación del movimiento y poder generar una cámara lenta de mayor calidad. Aunque suele ser a costa de bajar proporcionalmente la resolución espacial.

Para acelerar el movimiento por lo contrario tendremos que reducir el número de imágenes en la grabación. La técnica de *time-lapse* se utiliza para captar una acción que en la naturaleza se produce muy lentamente, por ejemplo el desplazamiento de las nubes. Para registrar este movimiento, se puede probar a grabar una imagen fija cada 10 segundos durante 42 minutos. El resultado será una secuencia de 250 imágenes que reproducidas a 25 fps, generarán una secuencia de 10 segundos en donde veremos fluir las nubes a gran velocidad.

Los sistemas de postproducción y composición digital gestionan la resolución temporal de los distintos materiales adaptando cada uno a las necesidades del proyecto y permiten editar dinámicamente la velocidad de reproducción. Como es lógico, las secuencias grabadas a una resolución temporal superior ocupan más espacio en el disco duro, puesto que están compuestas por un número mayor de imágenes fijas, y son más complejas de procesar.

La resolución temporal en la grabación también afectará a la definición de los bordes de los objetos en movimiento y al efecto de desenfoque de movimiento (*motion blur*). Un aumento en la frecuencia de captación implica también un aumento de la velocidad de obturación en la cámara generando así imágenes más nítidas con un efecto de desenfoque de movimiento reducido.

Algunos profesionales consideran que la descomposición del movimiento en 24 imágenes por segundo se percibe como 'efecto cine' mientras que las frecuencias más altas, tradicionalmente utilizadas en televisión, 50/60 campos por segundo, como 'efecto vídeo', más realista en la reproducción del movimiento.

5.4. Compresión de la señal de vídeo

5.4.1. Flujo de transferencia binaria

El flujo de transferencia binaria o tasa de transferencia (*bit rate*) indica el número de bits que se transmiten por segundo. Se puede expresar en bits por segundo (*bits per second*; bit/s) o en Bytes por segundo (B/s). Se suele utilizar la unidad bit cuando la transmisión se produce en serie, es decir, en fila uno detrás de otro y Bytes cuando la transmisión se hace con ocho bits en paralelo.

Cuanto mayor sea la resolución espacial, la profundidad de color y la resolución temporal, mayor será la cantidad de información (bits) a transmitir para poder reproducir el vídeo en tiempo real. Es decir, cuanta más calidad tenga el formato de vídeo, mayor será el flujo de transferencia que genera.

Por lo tanto, la tasa de transferencia es también un indicador de calidad técnica de la señal de vídeo. Por ejemplo, un códec con tasa de transferencia de 184 Mbps contiene más información que el de 36 Mbps y corresponde a una codificación de mayor calidad.

El cálculo del flujo de transferencia binaria es una operación matemática sencilla. Se trata de multiplicar el número de píxeles de resolución espacial por la profundidad de color y por la resolución temporal. Se obtiene de esta forma el número de bits que se han de transferir en un segundo para poder leer el vídeo en un dispositivo.

Experimentando esta fórmula con una resolución de 1920x1080, a 8 bits por canal y 25 fps, obtenemos los siguientes resultados:

- Resolución espacial: $1920 \times 1080 = 2.073.600$ píxeles por cada fotograma.
- Profundidad de color: $2.073.600 \times 8 \text{ bits} \times 3 \text{ canales RGB} = 49.766.400$ bits para cada fotograma.
- Resolución temporal: $25 \text{ fps} \times 49.766.400 = 1.244.160.000$ bit/s (bits por segundo).
- Flujo de transferencia binaria resultante: 1,2 Gb/s (Gigabits por segundo).

Este valor teórico de flujo binario, es demasiado alto y completamente inaceptable en la práctica para la radiodifusión o incluso difícil de gestionar para las aplicaciones profesionales más exigentes en términos de calidad. Se impone por lo tanto aplicar compresión y técnicas para optimizar el flujo binario o el ancho de banda de transmisión si lo expresamos con la terminología de la radiodifusión.

Como el ojo humano es más sensible al brillo que al color, para comprimir la señal de vídeo se puede hacer previamente una conversión de la señal RGB a la señal por componentes Y,U,V, que separa la luminancia de la crominancia. A partir de este primer paso se podrá aplicar mayor compresión al color que al brillo.

5.4.2. Señal de vídeo RGB y YUV

La señal de vídeo RGB utiliza un canal independiente para cada componente de color. Es el formato natural tanto de los equipos de captación como de reproducción. En este tipo de señal, la luminancia o brillo no se trata de forma independiente, sino que dependerá de la intensidad de cada uno de los colores primarios.

La separación de la información de luminancia y crominancia para usos profesionales se concretó en la llamada 'señal de vídeo por componentes' que es una adaptación de la señal RGB. Las tres señales de color se transforman en otras tres que representan, por un lado, la luminancia de la imagen, es decir la imagen en tonos grises, y que se representa por la letra Y; y por otro, se obtienen las dos señales que portan la información del color que se identifican con las letras U y V. También podemos referirnos a esta señal como Y, CB, CR.

Es decir que (Y) es la señal de luminancia y (U) y (V) son las componentes de crominancia.

La señal de luminancia (Y) está formada por un 30% de la señal roja (R), un 59% de la señal verde (G) y un 11% de la señal azul (B). La mezcla en estos porcentajes de rojo verde y azul da como resultado un pixel blanco. Son porcentajes diferentes por analogía a la percepción humana, que es más sensible a unos colores que a otros. Por lo que la luminancia (Y) o brillo queda expresada matemáticamente por la fórmula: $Y = 0,30R + 0,59G + 0,11B$.

Las señales que contienen la información de color (U y V), son la diferencia de las componentes rojo y azul con respecto de la luminancia. Se denominan señales de diferencia de color. La señal (U) es el resultado de restar el rojo con la luminancia (R-Y) y la señal (V), la resta del azul con la luminancia (B-Y). La componente verde se puede deducir de forma matemática.

En la siguiente tabla se resumen las ecuaciones que se utilizan para la conversión de la señal RGB a la señal por componentes YUV.

Tabla 5.10. Sistema de ecuaciones utilizada para la conversión de la señal RGB en vídeo por componentes Y,U,V

$Y = 0,30R + 0,59G + 0,11B$	Luminancia
$U = (R-Y)$	información del rojo menos la luminancia
$V = (B-Y)$	información del azul menos la luminancia

Fuente: *elaboración propia.*

Con esta fórmula se obtiene una señal de calidad equivalente a la RGB pero que separa la información de luminancia (Y) de la crominancia (U,V). Permitirá, entre otras aplicaciones, dar un tratamiento preferente a la luminancia en los procesos de codificación digital y compresión.

5.4.3. Submuestreo de color

El submuestreo de color (*chroma subsampling*) es un tratamiento de la señal de vídeo por el cual se reduce la información de color. Consiste en reducir el número de píxeles que se codifican en las componentes de color. Esta estrategia aprovecha que la visión humana es más sensible al brillo que al color. Así se conserva toda la resolución de la imagen en blanco y negro (brillo) y se elimina información de color, manteniendo la sensación de calidad.

Las normas de submuestreo de color que más se utilizan son las siguientes:

- Norma 4:2:2 (CCIR 601) Producción Calidad Estudio
- Norma 4:2:0 Periodismo electrónico y producción ligera (ENG)

Utilizando esta nomenclatura, 4:4:4 indica que no hay submuestreo de color y que por lo tanto se trata de una señal por componentes equivalente en calidad a RGB. Se utilizará en aplicaciones profesionales como por ejemplo la postproducción que requiere la información completa de luminancia y color de todos los píxeles para poder aplicar máscaras o efectos digitales. Las cámaras de cine digital y de televisión de alta calidad tampoco utilizan el submuestreo de color.

La norma 4:2:2 (CCIR 601) Producción Calidad Estudio, se definió para aplicaciones profesionales de televisión. De cada cuatro píxeles, se codificarán todos para la luminancia (Y) y tal solo dos para cada una de las señales de diferencia de color (U y V). Es decir que todos los píxeles llevan información de luminancia, pero solo uno de cada dos lleva información de color. La reducción en el flujo de transferencia binario es muy considerable y sin embargo la pérdida de información de color no es perceptible,

porque el ojo humano es más sensible a los matices y los cambios de luminosidad que a los de color (visión fotópica, mesópica y escotópica).

La norma 4:2:0 Periodismo electrónico y producción ligera (ENG), se utilizan en situaciones que requieren menor calidad como puede ser la producción de periodismo electrónico (programas informativos de TV) y las aplicaciones de vídeo doméstico. En este caso se codifica solamente la información de color de uno de cada cuatro píxeles: todos los píxeles activos tienen información de brillo pero solo una cuarta parte tienen toda la información de color. Aunque tenga menos registros de color que la norma 4:2:2, la percepción humana lo tolera sin que parezca un fallo cumpliéndose así con el objetivo de reducir la tasa de transferencia sin pérdida de calidad subjetiva. Sin embargo, esta norma no sería adecuada para su uso en aplicaciones de postproducción que requieren la información completa de color de cada píxel. Se utiliza habitualmente en las cámaras de bajo presupuesto y para la difusión de video por televisión e internet.

Con el submuestreo de color se consigue reducir considerablemente el flujo de transferencia binaria, aunque sea a costa de la resolución espacial del color. El objetivo es la reducción de la cantidad de datos a transmitir o grabar. Las técnicas de compresión de la señal también pueden contribuir eficientemente a cumplir este objetivo.

5.4.4. Compresión espacial y compresión temporal

La compresión implica una determinada forma de codificar la información que posteriormente tendrá que ser decodificada por el equipo de recepción. Cuando se codifica con una compresión con pérdidas, 'lossy', se elimina parte de la información que tendrá que ser reconstruida por el receptor.

Por ejemplo, si las 10 primeras líneas de una imagen son de color negro puro, no será necesario utilizar 24 bits para cada píxel en esa zona. Bastará utilizar un algoritmo que con unos pocos bits pueda expresar esa información con precisión. Se trata de transportar la misma información empleando menos datos.

Se diferencian dos tipos de compresión:

- La compresión espacial es la compresión que elimina la redundancia en el interior del espacio de una imagen (compresión intracuadro).
- La compresión temporal opera sobre el flujo redundante temporal (compresión intercuadro).

La compresión espacial trabaja únicamente con un fotograma (*frame*). Trabaja con la redundancia, o la similitud de los píxeles vecinos en la imagen, llamados bloques o macro bloques. Por ello se denomina compresión intracuadro (*intraframe*). Un ejemplo muy conocido de compresión espacial es JPG, que se utiliza de forma muy generalizada en fotografía y vídeo. Los códecs de vídeo DV, JPEG 2000 o los de la familia ProRes, en-

tre otros muchos, utilizan la compresión intracuadro. Este tipo de compresión se utiliza en los entornos de producción.

La compresión temporal compara la información de cuadros sucesivos para encontrar similitudes que se puedan expresar de forma más eficiente. Se denomina intercuadro (*interframe*) precisamente porque opera en una secuencia de imágenes en movimiento.

La compresión temporal funciona con paquetes de *frames* (*GOP, group of pictures*) que se codifican conjuntamente. Solo se guarda la información no redundante con respecto al primer fotograma del grupo que se establece como referencia. Este tipo de compresión se utiliza habitualmente en difusión.

Un problema que surge con la compresión temporal para la edición y postproducción de vídeo es que al hacer pausa en un fotograma concreto de la secuencia, se detiene el flujo y se pueden producir saltos, paradas inesperadas y falta de precisión. Para solucionar estos problemas los equipos tienen que utilizar mucha capacidad de cálculo ya que en cada parada tienen que descomprimir el GOP para reinterpretar el fotograma detenido.

Este tipo de compresión es muy eficiente para la transmisión de vídeo, pero problemática para la edición puesto que no existe una descripción completa de cada uno de los cuadros que componen la secuencia.

Los ejemplos más conocidos de compresión temporal son MPEG 2 y toda la familia de códecs basados en MPEG4 (H264).

5.5. Formatos de archivo y códecs

Los archivos digitales de imagen (*media files*) se pueden clasificar en tres grupos:

- 1) Los archivos RAW
- 2) Los formatos sin compresión (o con poca compresión)
- 3) Los formatos con compresión

Los archivos RAW contienen toda la información en bruto que ha captado el sensor de la cámara. No se pueden ver de forma inmediata ya que falta aún un realizar el procesado de la imagen. Cuando intentamos ver en un ordenador un archivo RAW se abre un programa de 'revelado RAW'² en donde tendremos que definir los parámetros básicos, como por ejemplo la ganancia, el balance de blancos o la profundidad de color, con la que queremos completar la codificación de la imagen. Por lo tanto implica un proceso de postproducción. De ahí su nombre que en inglés significa 'crudo'. Es el formato que mayores posibilidades ofrece en términos de calidad técnica para el retoque fotográfico y la postproducción de vídeo.

² Ref. artículo sobre revelado RAW y profundidad de color: <http://todo-fotografia.com/2013/profundidad-de-color/>

Las imágenes codificadas sin compresión, o con poca compresión, tienen la ventaja de preservar la calidad original del material y el inconveniente de que requieren mucho espacio en el disco duro para su almacenamiento. En otras palabras ‘pesan mucho’.

Los archivos de imagen con compresión tienen la virtud de ocupar poco espacio en el disco duro o en las tarjetas de memoria de las cámaras. El inconveniente es que no preservan toda la calidad posible como los archivos RAW o los que no tienen compresión. Se denomina ‘compresión sin pérdidas subjetivas’ a la compresión que no es perceptible a la vista en condiciones normales de uso. Pero aun así, con los sistemas de compresión, hay mucha información que se pierde en la codificación. En inglés se utilizan los términos ‘*loosy compression*’ y ‘*lossless compression*’ para referirse a este concepto.

5.5.1. Códecs, contenedores y formatos de vídeo

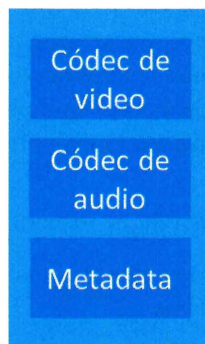
Las distintas pistas de vídeo, audio y metadatos que componen un fichero de vídeo se compilan en formatos contenedores (*container formats*).

Como ejemplos de formatos contenedores muy conocidos se pueden citar AVI, que es el contenedor estándar de los equipos de Microsoft, MOV el de los ordenadores de Apple o MP4 que es el contenedor que se utiliza para los archivos de vídeos codificados en MPEG-4. En entornos profesionales de producción se utiliza frecuentemente el formato contenedor MXF (*Material Exchange Format*).

Las especificaciones técnicas de la codificación de cada pista de vídeo y audio se denominan ‘códec’. Este término proviene de la contracción de ‘codificación y decodificación’.

Resumiendo, un archivo de vídeo tiene un formato contenedor y un códec. Y dentro de un formato contenedor puede haber distintos códecs.

Figura 5.11. El formato contenedor encapsula en un solo archivo el vídeo, el audio y los metadatos



Fuente: elaboración propia

Para que un archivo de vídeo se pueda leer en un ordenador es necesario instalar previamente el códec, donde se describen los algoritmos de codificación y compresión que se han utilizado al generarlo.

En la siguiente figura se puede ver el menú de exportación del programa 'DaVinci Resolve' con desplegados para elegir el formato contenedor y el códec de vídeo.

Figura 5.12. Menú de exportación en 'DaVinci Resolve' con desplegados para seleccionar el formato contenedor (MXF) y el códec (DNxHR)



Fuente: *elaboración propia*

Se pueden clasificar los códecs en cuatro grupos:

- Códecs nativos de cámara
- Códecs de intermediación o de postproducción
- Códecs de masterización y archivo
- Códecs de difusión

Pero estos grupos no son estancos, puesto que hay cámaras que ofrecen la posibilidad de grabar con códecs que estarían naturalmente en el apartado de postproducción (por ejemplo ProRes) o en el de difusión (por ejemplo H.264). Lo mismo ocurre con los sistemas de edición y postproducción que pueden trabajar con cualquier códec.

Códecs nativos de cámara

Se denominan 'códecs nativos de cámara' a los sistemas de codificación que utilizan los distintos fabricantes de cámaras.

Cada fabricante de cámaras de cine digital utiliza su propio códec. Por ejemplo, la cámara 'Arri Alexa', considerada actualmente como una de las mejores cámaras de

cine, utiliza un códec propietario que se llama Arri RAW³. En cine digital la tendencia es grabar en formato RAW o en imagen procesada aplicando una 'curva de *gamma* logarítmica'. Será necesario por lo tanto pasar por un proceso de postproducción y corrección de color para ajustar el contraste y aprovechar todo el rango dinámico registrado.

Las cámaras orientadas al mercado de la televisión permiten configurar la grabación para exploración entrelazada o progresiva. La imagen que registran está ya adaptada a la norma técnica de radiodifusión (Rec-709), es decir imagen procesada con una 'curva de *gamma* corregida'.

Las cámaras ligeras y de bajo presupuesto como las réflex DSLR suelen grabar en tarjetas SD, utilizan códecs con compresión *lossy* (por ejemplo h264), con submuestreo de color 4:2:0 y con una profundidad de color de 8, 10 o 12 bits. Generalmente ofrecen la opción de elegir el formato de contenedor Quick Time (.mov) o MP4 (.mp4). Algunos modelos ofrecen la posibilidad de grabación en formato RAW, pero generalmente registran imagen procesada con ajustes de *gamma* Rec-709.

Códecs de intermediación o de postproducción

Los códecs de intermediación se emplean en postproducción de alta calidad, para el montaje *online* / *offline* y para la corrección de color. Tratan de mantener la calidad de las imágenes originales con una compresión mínima y procurando que el peso final de los archivos no sea muy alto. En el caso del montaje *offline*, se usarán formatos con mayores niveles de compresión.

Utilizan solo compresión espacial *intraframe*, es decir comprimen independientemente cada fotograma de la secuencia.

Existen muchos códecs de intermediación, aunque los más utilizados son los ProRes de Apple y los DNxDH de Avid. En los sistemas de postproducción sobre la plataforma Mac se suele utilizar los ProRes. Sobre PC, es menos frecuente ya que Apple cobra una licencia para permitir codificar en ProRes. Sin embargo la familia de códecs de Avid, DNxHD, se distribuye gratuitamente en internet, tanto para PC como para Mac y es equivalente en calidad.

Códecs de masterización y archivo

Una vez terminada la postproducción se utilizará un códec sin compresión (o con poca compresión) para archivar el material. En este momento el espacio que ocupe el fichero ya no es tan relevante. Se suelen utilizar, tanto para la masterización como para archivo, códecs sin compresión o con compresión sin pérdidas. Generalmente se

³ La cámara Arri Alexa, también ofrece la posibilidad de grabar con el códec ProRes de Apple.

graba cada fotograma en una imagen fija independiente, por ejemplo en formato TIFF. La ventaja de guardar una secuencia de imágenes fijas es que si por algún motivo se corrompe un archivo, el resto de los fotogramas quedan intactos. Ofrece también cierta garantía de que el formato no quedará anticuado o descatalogado con el paso del tiempo. El inconveniente de estos formatos de archivo es que ocupan muchísimo espacio de almacenamiento.

Los formatos de archivo más utilizados son DPX, QT sin compresión y MXF. El formato DPX es más recomendable hoy en día porque MXF o QT demandan más capacidad de proceso a la máquina.

Códecs de difusión

El estándar de difusión del cine digital es el DCP (*Digital Cinema Package*) (ver capítulo 12). El DCP utiliza el códec JPEG 2000, que es una secuencia de fotogramas con compresión JPG que se encapsula en MXF. Tiene un sistema propio de encriptación para evitar usos ilegítimos de las copias de alta calidad que se proyectan en las salas.

En radiodifusión de televisión se utilizan códecs orientados a optimizar el ancho de banda de transmisión. Se utilizan los códecs MPEG2 o MPEG4 (H.264) para la televisión digital de resolución estándar y de alta definición. El nuevo salto en esta familia de códecs es el MPEG-H Vídeo, más conocido como HEVC (H.265), el códec que se está utilizando para ultra alta definición.

Para la difusión por internet se utilizan ficheros muy ligeros con mucha compresión. Los formatos H.264, basados en la especificación MPEG4, ofrecen toda una gama de *preset* optimizados para la difusión por internet. Para la difusión de contenidos audiovisuales de alta calidad por internet existe la norma MPEG DASH (*Dynamic Adaptive Streaming over HTTP*) que recomienda, para el *streaming* 4K, el encapsulado MPEG y, entre otros, el códec HEVC (H.265).

5.5.2. Comparación de dos formatos de video

En la siguiente tabla se comparan dos formatos de video: un formato de difusión televisiva de alta definición que se utiliza en Europa, [DVBT 1080i, 25 fps] y un códec de alta calidad de los que se utilizan en postproducción profesional, [DNxHD 365x, 1080p, 25 fps].

Tabla 5.13. Comparación de formatos de archivo de alta definición

	DVBT 1080i/25	DNxHD 365x,1080p/25
Resolución espacial	1920x1080	1920x1080
Relación de aspecto	16:9	16:9
Profundidad de color	8 bits	10 bits
Proporción del píxel	Cuadrado	Cuadrado
Submuestro de color	4:2:2	4:4:4
Resolución temporal	25 fps / 50 campos	25 fps
Tipo de exploración	Entrelazada	progresiva
Compresión	MPG4	Sin compresión
Flujo binario	variable entre 10 y 4 Mb/s	367 Mb/s

Fuente: elaboración propia

Se puede observar que la resolución espacial de ambos formatos es la misma, sin embargo el flujo binario que genera la codificación pasa de los 367 Mb/s del códec sin comprimir a una cifra variable que puede ir entorno a los 5-10 Mb/s para la difusión de TDT.

Un flujo de transferencia de 365 Mb/s solo lo soportará una estación de trabajo profesional de postproducción de cine digital y televisión. Con este nivel de calidad en un gigabyte de memoria en el disco duro solo se podrán almacenar 23 segundos de video. Para un ordenador convencional será difícil reproducir este archivo.

El códec de DVB tiene la compresión *interframe* del MPEG 4 y submuestreo de color 4:2:2. Es muy eficiente para la radiodifusión televisiva, pero no es adecuado para su uso en postproducción profesional.

Referencias bibliográficas:

- Armenteros, M. & Utray, F. (2011). "La señal de vídeo". En Armenteros, M. (Dir.), *Posproducción Digital*. Bubok.
- Brinkmann, Ron (2008), *The art and science of digital composing* (Third edition). Morgan Kaufman.
- Bethencourt, Tomás (2003), *Televisión Digital* (2 ed.), I. G. Afanias, Colección Beta, Temas audiovisuales.
- Bordwell, D. y Thompson, K., (1995), *El Arte Cinematográfico*, Paidós.
- Carrillo de Albornoz, Ricardo (2013) "Profundidad de color" artículo publicado en la web Todo-fotografía, disponible para descarga en [<http://todo-fotografia.com/2013/profundidad-de-color/>]
- Castillo, José María, (2011), *Televisión, realización y lenguaje audiovisual*, cap. 6 Tecnología de la televisión, Instituto de Radio y Televisión Española.
- Ochoa, Luis (2014), *Apuntes del curso de corrección de color para cine y televisión*, Inédito, curso impartido en junio 2014 en la Universidad Carlos III de Madrid.

Lecturas recomendadas:

- Avid DNxHD Technology. Beauty without the bandwidth. Disponible para descarga en <http://www.avid.com/static/resources/es/documents/DNxHD.pdf>
- ProRes 2K in Editorial White Paper. Disponible para descarga en <http://www.ari.com/downloads/>