

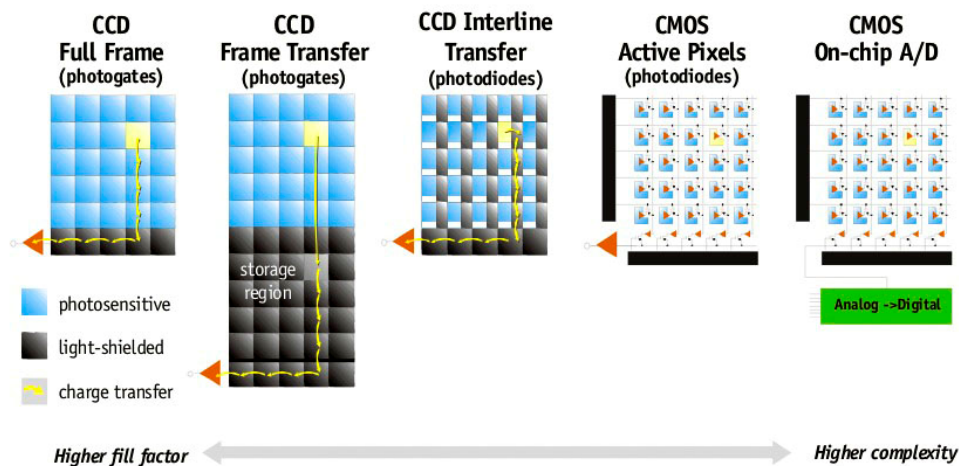
Fundamentos de fotografía digital

Efraín García y Rubén Osuna ©

8. Variedad tonal y amplitud tonal.

La calidad potencial de una imagen fotográfica digital no sólo depende de variables como la resolución y el contraste. La *variedad tonal* y la *amplitud tonal* son también variables determinantes para la calidad de la imagen fotográfica y, aunque hemos ido apuntando su importancia, es verdad que nos hemos concentrado mucho en la pura capacidad de resolver detalle de las cámaras.

Hablamos de *espaciado* entre ‘píxeles’ porque el tamaño real de las celdillas que capturan luz depende de la superficie del sensor que queda ocupada por dispositivos –circuitos– que procesan la señal luminosa y transportan la carga eléctrica resultante hacia los bordes del mismo. El término *fill factor* alude precisamente a la *proporción efectivamente sensible a la luz del sensor*.



Distintos tipos de sensores CMOS (Metal Oxide Semiconductor) y CCD (Charge Coupled Device) ([Dalsa©](#)).

La proporción entre señal falsa (ruido) y señal real mejora no sólo reduciendo las causas del ruido (->), sino también mejorando la eficiencia cuántica (->) y aumentando la cantidad de señal por ‘píxel’ mediante mejores filtros, microlentes y un *fill factor* mayor (celdas más grandes). Es por tanto lógico que, en efecto, se hayan observado mejoras en la razón señal-ruido (*signal-to-noise ratio*, SNR), incluso si el *espaciado* entre los ‘píxeles’ disminuye. Estos avances en la SNR han permitido, entre otras cosas, mejorar la “resolución tonal” (o “variedad tonal”) y la “amplitud tonal” (o “rango tonal”, “escala tonal”, “intervalo tonal”) de las cámaras digitales, a la vez que se incrementaba la resolución. Este progreso tiene, lógicamente, un límite, pues hay restricciones puramente físicas que no se pueden burlar.

Tradicionalmente han sido los sensores CCD de tipo “full-frame” (o los “frame-transfer”, muy similares, empleados en cámaras de cine digitales) los que ofrecían mayor superficie sensible a la luz (con un *fill factor* muy cercano al 100%), si bien los sensores CCD de tipo “interline transfer” y los CMOS han experimentado grandes mejoras en este apartado (los CMOS tienen un *fill factor* típico de en torno al 60%, con ayuda de microlentes). El progreso

de estos dos tipos de sensores, pero especialmente el de los CMOS, está condicionado por los avances en la tecnología de litografía aplicada en su producción, que determina el tamaño de los circuitos, pero también por límites físicos: una reducción del tamaño de los circuitos plantea problemas de recolección y medición de la carga eléctrica generada por las fotocélulas (->). Los CMOS, con más dificultades para aumentar la superficie de captura, se han especializado en sofisticados tratamientos electrónicos de reducción del ruido (varios transistores en cada píxel, lecturas de señal en el propio sensor), lo que unido a su menor consumo de energía y voltaje de operación, ha permitido conseguir notables rendimientos en cuanto a razón señal-ruido, especialmente cuando reducimos la señal (subimos la sensibilidad equivalente ISO del sensor). Los CCDs, por contra, siguen proporcionando la mejor relación señal-ruido para valores ISOs bajos.

La superficie perdida con circuitos puede tratar de recuperarse mediante el uso de *microlentes*. Las microlentes son microscópicas lupas que corrigen el ángulo de incidencia de la luz, concentrándola en un punto. Permiten recoger luz que caería en las zonas "insensibles" del sensor y redirigirla hacia las zonas sensibles. Además ayudan a resolver otro problema: entre la superficie del sensor y el lugar que ocupa el fotodiodo –que recibe luz y genera una carga eléctrica– hay una distancia (el problema es más serio en los CMOS debido a la presencia de 3, 4 o incluso 5 transistores en cada 'píxel'). Por decirlo de algún modo, los fotodiodos están en el fondo de un pozo tapado por gruesas capas de filtros translúcidos. La consecuencia es que si la luz incide con un ángulo demasiado inclinado acabará desviándose, perdiéndose entre los filtros o chocando con la pared lateral del pozo. Una forma de aliviar este problema es reducir la gruesa capa de filtros. Pero otra es concentrar y dirigir la luz hacia el fotodiodo.

El problema de las microlentes es que no se adaptan al ángulo de entrada de la luz cuando éste varía, pues son fijas. La abertura y la longitud focal, entre otras cosas, modifican aquel ángulo, por lo que microlentes eficientes bajo ciertas condiciones pueden dejar de serlo en otras. Pero la tecnología de microlentes también ha mejorado, permitiendo una reducción del espacio entre ellas y modificando su orientación en función de la posición que ocupan en el sensor (el caso del sensor [CCD Kodak](#) de la [Leica M8](#)).

La *amplitud tonal* es la razón entre el máximo nivel de luminosidad que el sensor puede medir antes de saturarse y el mínimo nivel descontado el ruido de lectura (->). Fuera de ese rango la cámara percibe un negro o un blanco absolutos. Depende, fundamentalmente, del tamaño real de las celdillas en que se alojan las fotocélulas y de la capacidad de diferenciar matices del conversor analógico-digital. Esta *amplitud tonal* se mide, en términos fotográficos, en "pasos" (*stops*), en una escala logarítmica en la que cada "paso" implica doblar la cantidad de luz del paso anterior ($2 \times 2 \times 2 \dots \times 2 = 2^n$ donde n es el número total de pasos). Las mejores cámaras digitales tienen una escala tonal de unos 8 ó 9 pasos (->), y excepcionalmente hasta 10 pasos (->), mientras que los negativos químicos se mueven dentro de un rango similar y las diapositivas en uno más limitado (->). Con rápidos ajustes de la pupila el ojo humano puede abarcar hasta 24 pasos de escala tonal, y con una abertura constante entre 10 y 14 pasos.

Surge un problema importante debido al carácter *lineal* del soporte digital que almacena la información fotográfica (los famosos 12, 14 o 16 bits de variantes tonales por 'píxel'), frente al carácter *logarítmico* de la información luminosa.

Si colocamos una vela delante de una cámara un determinado tiempo, cada 'píxel' del sensor recogerá una determinada cantidad de luz y generará una carga eléctrica. Si colocamos dos velas, la carga eléctrica será doble en el mismo tiempo. Triple si son tres velas. Etcétera. La

escala es lineal (1, 2, 3, 4...). Sin embargo, el ojo humano percibe una duplicación de la intensidad luminosa si pasamos de una vela a dos, pero no si pasamos de dos velas a tres. El ojo (o el cerebro) mide las diferencias logarítmicamente, en una escala en la que cada valor duplica al anterior (1, 2, 4, 8, 16, 32...). Con este ajuste logarítmico nos protegemos de una posible sobredosis luminosa (al pasar de 16 velas a 32 el efecto percibido es similar al de pasar de 1 a 2), lo que nos permite percibir una mayor amplitud tonal. La película responde a la luz de forma similar al ojo humano, es decir, no linealmente, y esta es una de las grandes diferencias entre el medio químico y el digital (el único sensor que se aproxima a esta respuesta no lineal es el SuperCCD de Fujifilm).

Cada fotocélula transforma el estímulo luminoso en una carga eléctrica, es decir, responde al impacto de los fotones generando electrones. Los CCDs transportan esta carga eléctrica secuencialmente hacia los bordes del sensor, de 'píxel' a 'píxel', y un colector la transforma en un determinado voltaje. Los CMOS hacen la transformación de la carga eléctrica en voltaje en el propio 'píxel' y la transportan en paralelo (de forma más rápida). El conversor analógico-digital (*A/D converter*) hace una lectura de esos voltajes y genera una secuencia de números digitales, siguiendo una escala lineal (\rightarrow). Para pasar de una escala lineal a una logarítmica:

$$\text{Pasos} = \log_2(\text{Medida lineal}) = \ln(\text{Medida lineal}) / \ln(2)$$

Donde la medida lineal aparece siempre que expresamos la magnitud del contraste (por ejemplo, 10.000:1, donde 10.000 sería la medida lineal).

Sin embargo, al superponer una escala logarítmica a la escala lineal, el paso más alto de luz se queda con el 50% de esas 2^{12} , 2^{14} ó 2^{16} variantes tonales en las que el conversor analógico-digital "traduce" su lectura de la carga eléctrica (los famosos 12, 14 o 16 bits). El siguiente paso, más bajo, se queda con la mitad del sobrante ($0,5 \cdot 0,5 = 0,25$ o el 25% de las tonalidades alternativas), el siguiente paso más bajo se queda con el 50% del sobrante ($0,5 \cdot 0,5 \cdot 0,5 = 0,125$ que es el 12,5%), etc. Es obvio que los pasos más bajos apenas se quedan con variantes que describan sus matices (véase la siguiente tabla). Se puede observar cómo 8 bits de variantes tonales limitan la máxima amplitud tonal teóricamente posible (sin considerar las limitaciones del ruido) a 8 pasos, 12 bits a 12 pasos, 14 bits a 14 pasos, etc. A partir de ese punto, las variantes tonales se agotan. Por tanto, la precisión del conversor analógico-digital limita la amplitud tonal máxima teóricamente reproducible. Hay que tener en cuenta además que el ruido inutiliza o degrada apreciablemente un buen número de variantes tonales en las sombras, ya que ahí la señal por 'píxel' es lógicamente pequeña. Por ello la amplitud tonal es siempre inferior al máximo que teóricamente puede manejar el conversor analógico-digital (una regla orientativa es restar al menos dos pasos).

En general, los sensores CCDs de tipo "full-frame" y "frame-transfer" tienen una razón señal-ruido más alta que los CMOS, y eso les permite "salvar" más pasos de luz y ofrecer una mayor amplitud tonal. Una forma muy común de reducir el ruido que el usuario ve en la imagen final obtenida consiste en, simplemente, sacrificar amplitud tonal, descartando las zonas de sombras altamente "contaminadas". Una conclusión interesante es que *una mayor amplitud tonal tiene que estar acompañada de una mayor profundidad de bits, aunque ésta no garantiza aquélla*, debido en parte al ruido. Por otro lado, la reducción del ruido mediante avances en electrónica permite "rescatar" información en las sombras y ganar amplitud tonal. Esto es lo que hace Canon en sus CMOS, mediante la incorporación transistores especializados en cada 'píxel' que consiguen reducir una de las fuentes de ruido (el 'reset noise'), aunque por otro lado afectan al *fill factor*.

La amplitud tonal es como la longitud total de una escalera, donde la profundidad de bits viene a ser el número de escalones. No obstante, la altura de estos escalones no es la misma en todo el recorrido. En la parte baja de la escalera (sombras) los escalones son escasos en número y muy altos, y conforme vamos subiendo (luces altas) estos se hacen más numerosos y más bajos. Si queremos una escalera más alta tendremos que añadir escalones. Pero no es posible incrementarlos sólo en la parte baja, donde se agradecería un mayor número de ellos. Los nuevos escalones tienen que distribuirse desigualmente a lo largo de toda la escalera, acumulándose preferentemente en la parte alta, ¡donde menos falta hacen!

%	paso	Lineales				No lineales	
		8 bits, $2^8 = 256$	12 bits, $2^{12} = 4.096$	14 bits, $2^{14} = 16.384$	16 bits, $2^{16} = 65.536$	16bits, Gamma 2.2	Leica DNG 8 bits
		variantes tonales	variantes tonales	variantes tonales	variantes tonales	variantes tonales	variantes tonales
50,0	n	128	2.048	8.192	32.768	17.712	74
25,0	n-1	64	1.024	4.096	16.384	12.925	54
12,5	n-2	32	512	2.048	8.192	9.415	37
6,25	n-3	16	256	1.024	4.096	6.883	27
3,13	n-4	8	128	512	2.048	5.023	18
1,57	n-5	4	64	256	1.024	3.665	14
0,78	n-6	2	32	128	512	2.675	9
0,39	n-7	1	16	64	256	1.952	7
0,20	n-8	-	8	32	128	1.425	4
0,10	n-9	-	4	16	64	1.039	4
0,05	n-10	-	2	8	32	758	7

Variantes tonales y su distribución por pasos. Las zonas sombreadas están más afectadas por el ruido.

Debe distinguirse entre las variantes que maneja el convertor analógico-digital de las empleadas finalmente para el “almacenaje” de la información en el archivo final, que pueden comprimirse a 8 bits por canal o quedar sin comprimir a 16 bits. Sin embargo, también es posible aplicar un algoritmo de compresión no lineal que reduzca el número de “escalones”, uniendo muchos de los pequeños (que captan pequeñísimas variaciones allí donde el ojo humano no diferencia muy bien, en las luces altas) y dejando intactos los grandes. Es lo que hace Leica con los archivos DNG de la M8 (a 256 variantes tonales) o Nikon con los NEF (a 683 variantes). Veamos a título de ejemplo cómo redistribuye la información luminosa la Leica M8. El convertor analógico-digital es de 16-bits, pero se descartan dos pasos de información, afectada por el ruido. El *output* es por tanto de 14-bits. Estas 16.384 variantes tonales *posibles* (no necesariamente hay información relevante para todas) se reducen a 255 empleando un algoritmo sencillo: el valor registrado en cada píxel se multiplica por 4 y al resultante se le calcula la raíz cuadrada [$\text{SQRT}((14\text{-bit data})^4)$]. En verdad, más que calcular raíces cuadradas, se aplica una tabla de correspondencias a la matriz de datos original, conocida como *Lookup Table* (incluyéndose la inversa en el archivo DNG). Esto reduce las 16.384 variantes originales a 255 (->). En el fondo estamos aplicando una sencilla curva de contraste ($y = 2x^{1/2}$, donde y son los valores transformados y x los valores originales, que es aproximadamente una gama 2,0), que preserva las variantes tonales en las zonas de sombra (o las reduce muy poco) y reduce el número de variantes reservadas a las luces altas, donde el ojo humano apenas percibe diferencias (el planteamiento de Nikon

es similar, aunque más holgado, \rightarrow , \rightarrow). El número de variantes reservado a cada paso de luz por la Leica M8 se muestra en la tabla anterior.

Esta compresión no lineal sólo es posible en el almacenamiento de la información. Todos los programas informáticos trabajan linealmente, por lo que al “leer” el archivo hay que pasar de nuevo a un espacio lineal donde, una vez más, las luces altas tendrán reservada en principio una cantidad desproporcionada del total de variantes tonales posibles. En la Leica M8 el algoritmo de descompresión es justo el inverso (algo así como dividir por 4 y elevar al cuadrado, mediante la aplicación de la tabla inversa mencionada). Es obvio que no se obtiene como resultado la diferenciación tonal original, y que el proceso implica pérdidas de información. Sin embargo, la distribución de variantes tonales entre los distintos pasos de luz es mucho más eficiente que en los 8 bits lineales de los archivos JPG y TIFF, por lo que no son comparables.

Durante el tratamiento de la imagen es conveniente trabajar *siempre* en el mayor espacio posible, de 16 bits o superior. Podemos pensar que cada variante tonal que es capaz de separar un conversor A/D es un “cajón”, y la información luminosa captada serían “documentos”. Cuando alteramos una imagen digital, tocando los niveles o la curva de contraste, movemos esos documentos de un cajón a otro, uniéndolos a los que allí se encontraban originalmente. Una vez hecho esto ya no hay vuelta atrás. No hay manera de saber de qué cajón vienen originalmente los documentos que nos encontramos en cualquiera de ellos en un momento dado. Cuantos más movimientos de documentos hacemos, más tienden a mezclarse y agruparse éstos en unos pocos cajones, especialmente en las luces bajas. Esto provoca una acumulación de información indiferenciada dentro de una estrecha escala tonal, que se muestra como desagradables empastes o apelmazamientos en la imagen (*posterización*). No hay que olvidar que cuando se “lee” un archivo RAW *siempre* se aplica una curva de contraste que corrige la distribución excesivamente desigual de variantes tonales entre pasos. Hemos mostrado en la tabla el resultado de un ajuste a una gama 2,2 (\rightarrow), donde $VTR = VT^{(1/G)}$, siendo VTR las variantes tonales resultantes, VT las variantes tonales originales y G la gama.

[Efraín García](#) y [Rubén Osuna](#) ©