

LA FOTOGRAFÍA Y LA DIGITALIZACIÓN MASIVA DE DOCUMENTOS



Seminario
organizado por la asociación

fotoConnexió

con la colaboración de



Impartido por
Esteve Cruanyes y
Francesc Fabregat
directores de



LA FOTOGRAFIA Y LA DIGITALIZACIÓN MASIVA DE DOCUMENTOS

Tabla de contenido

1.- PRESENTACIÓN DE LA DIGITALIZACIÓN MASIVA.....	3
1.1.- La digitalización del fondo SACE.....	4
del “Institut Cartogràfic de Catalunya” (ICC)	4
1.2.- La digitalización del fondo gráfico del Grupo Godó	4
- La Vanguardia y el Mundo deportivo -	4
1.3.- La digitalización del fondo Casañas	5
- “Arxiu històric de Sabadell” -	5
1.4.- La digitalización de fondos fotográficos del.....	5
- “Arxiu Nacional de Catalunya” -	5
2.- DIFERENCIAS CONCEPTUALES ENTRE LA FOTOGRAFÍA CLÁSICA Y LA DIGITALIZACIÓN MASIVA DE DOCUMENTOS	9
3.- LOS SENSORES FOVEON v.s. MOSAICS.....	11
4.- EL PROCESO DE “DEMOSAICING”.	15
IMPLICACIONES EN EL CONCEPTO DE RESOLUCIÓN.....	15
4.1.- El proceso de “demosaijing”	15
4.2.- Implicaciones en el concepto de resolución.....	16
5.- LA RESOLUCIÓN. UN EJEMPLO DE DISONANCIAS ENTRE FOTÓGRAFOS, IMPRESORES Y DIGITALIZADORES	17
5.1.- La resolución según los digitalizadores	18
5.2.- La resolución según los fotógrafos.....	18
5.3.- La resolución según los impresores	19
6.- EL BAYER, EL RAW Y EL TIFF	21
6.1.- Algunos ejemplos de retoques dinámicos automatizados	22
7.- LAS NUEVAS CÁMARAS DE 71 MP EN LA DIGITALIZACIÓN DE DOCUMENTOS.....	28

1.- PRESENTACIÓN DE LA DIGITALIZACIÓN MASIVA

■ ¿Qué digitalizamos?

- Papel “suelto” (suelto, grapado, clipado, más o menos cosido, etc.)
 - Expedientes administrativos, Expedientes de obras
 - Legajos , Carpetillas ...
 De todos los tamaños y tipologías
 - Mapas, acetatos, planos, carteles, cuadros,...
 - Pergaminos, “acordeones”, rollos, ...
 - Legajos (“muy” cosidos (recosidos varias veces)) ...
 Fotografías de todos los tamaños y “colores”
- Papel encuadernado
 - libros, revistas, prensa encuadernada, álbumes, cantorales, cartularios, ...
- Negativos y transparencias de todos los tipos y tamaños
 - Negativos de 35mm, de 6x6, 6x9, ..., 9x12, 10x13, ...
 - Acetatos, poliésteres y transparencias hasta el DinA3
 - Diapositivas
 - Vidrios hasta el DinA3, Pares estereoscópicos,...
- Microfilms
 - Rollos de 35 mm, de 16 mm
 - Microfichas. Tarjetas de apertura...
- ...

Pero siempre en cantidades muy grandes: **de decenas de miles a millones** de imágenes en cada proyecto.

Entregándose en general:

- una copia máster en formato TIFF sin compresión
- a veces también una copia en RAW tal cual sale de la cámara y en formato **propietario** del constructor de la cámara
- una copia en formato JPEG “optimizada” para su difusión en Internet
- si hay una unidad superior a la imagen individual, como sería el caso de un legajo , revista, expediente, etc. se acostumbra a entregar un PDF en formato PDF/A2 y si el contenido es básicamente texto de tipo mecanografiado, el PDF debe incorporar una capa de texto reconocido mediante OCR (alguna vez nos han pedido el PDF con índice de marcadores que generamos automáticamente)
- ...

Por lo que respecta a los metadatos, éstos acostumbran a componerse de:

- Datos automáticos del procedimiento de digitalización
- Datos preexistentes en la base de datos del archivo
- Datos documentados por el operador de digitalización
- Observaciones e incidencias ...

Los cuales se entregan

- Unos incrustados en las propias imágenes (/PDF)
- Otros en los diversos formatos “oficiales” para ese tipo de datos:

1.1.- La digitalización del fondo SACE del “Institut Cartogràfic de Catalunya” (ICC) <http://cartotecadigital.icc.cat/cdm/landingpage/collection/fsace>

Siguiendo el hipervínculo, podremos consultar las imágenes que digitalizamos de dicho fondo, constituido por unos 23.700 negativos de diversas medidas no estándar, aunque en general eran de 13x13 cm, correspondientes a imágenes aéreas oblicuas de Catalunya en no muy buen estado.

Digitalización en color de 24 bpp con una resolución de 2032 ppp sobre el original

Las principales dificultades fueron las siguientes:

- El formato no estándar de los negativos
Que obligó a construir soportes especiales que pudieran adaptarse a la medida de cada negativo en particular y que evitase el contacto de la película con el vidrio del escáner para que no se formaran anillos de Newton.
- El recorte de los bordes del negativo, a fin de entregar una imagen limpia al ICC

El problema radicaba en que la imagen ocupaba a menudo la zona del borde, aunque en él quedaba bastante difuminada. Así, había bordes de todo tipo: negros (caso ideal), grisáceos con imagen corrida, oscuros con muchos niveles de gris presentes y muy parecidos a los niveles de gris de la imagen cercana a los bordes, trozos oscuros del ala del avión metidos en la imagen, etc. Las anchuras de los bordes variaban de un negativo a otro, así como de un lado a otro en un mismo negativo.

Diseñar un algoritmo automático capaz de recortar correctamente los negativos fue muy complejo.

1.2.- La digitalización del fondo gráfico del Grupo Godó - La Vanguardia y el Mundo deportivo - https://www.upf.edu/hipertextnet/numero-2/gestion_imagen.html



Se digitalizaron los fondos gráficos del Mundo Deportivo y de La Vanguardia en sus respectivas sedes de Barcelona.

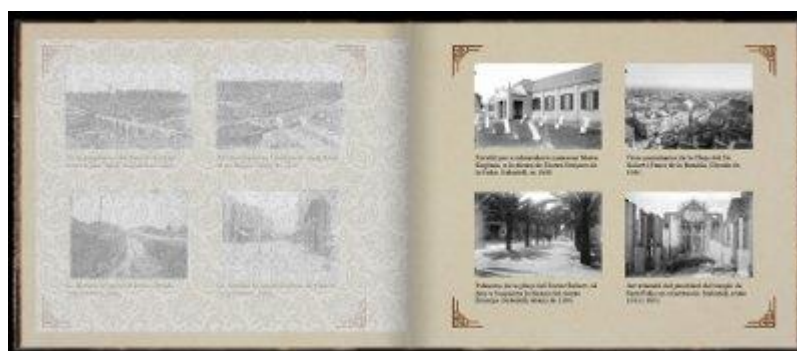
Los fondos, muy extensos, de varios millones de imágenes digitalizadas, y compuestos por una gran variedad de tipologías: fotografías papel, diapositivas, negativos en tiras, vidrios, carteles, etc.

El Objetivo era dotar, a los periodistas del grupo, de acceso al fondo gráfico para poder usarlo en su trabajo diario, como fuente de ilustraciones o meras consultas. Para ello debía documentarse en cuanto a personajes, situaciones, localizaciones, fechas, etc., de forma que los periodistas pudieran buscar y encontrar documentos relacionados con su interés en cada momento. Esta necesidad de completar al máximo los originales fue uno de los problemas a resolver.

1.3.- La digitalización del fondo Casañas

- “Arxiu històric de Sabadell” -

Para el “Arxiu Històric de Sabadell” se realizaron distintos proyectos de digitalización de fondos fotográficos, entre los cuales señalaremos el del fotógrafo Francesc Casañas tanto por la importancia y lo peculiar del personaje, como por el hecho que se montó un **álbum virtual** que fue de los primeros que se montaron en España, y posiblemente continúe siendo uno de los más completos en cuanto a prestaciones, aunque se desarrolló con tecnología Silverlight, hoy ya un tanto obsoleta



<http://ca.sabadell.cat/albumfotos/>

1.4.- La digitalización de fondos fotográficos del

- “Arxiu Nacional de Catalunya” -

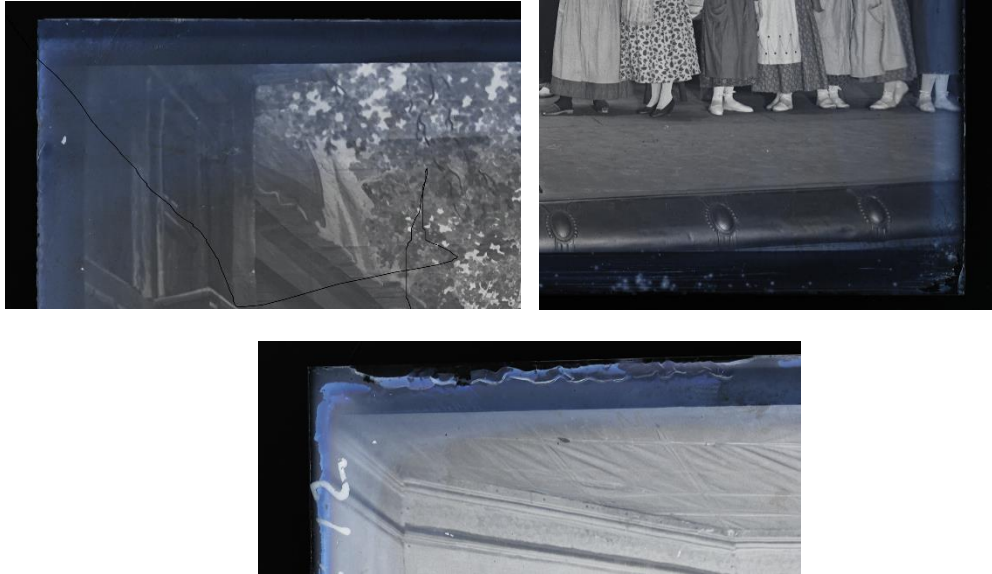
Se han realizado varios proyectos de digitalización para el “Arxiu Nacional de Catalunya” de los cuales referenciamos los que, por una u otra razón, son los más curiosos,

Fondo Brangulí

Fondo constituido negativos sobre vidrio de unas 35.000 imágenes, que debían entregarse con una profundidad de color de 8 bpp y a una resolución variable según el tamaño del original (la resolución se adecuaba a las medidas del original porque el cliente quería que al imprimir la imagen a 300 ppp le ocupase un DinA3)

Se digitalizó por transparencia sobre una mesa de luz de leds i utilizando una cámara Nikon de 36 Mp (D800E). La digitalización se realizaba en color de 24 bpp y posteriormente se pasaba a grises de 8 bpp.

El problema era que los bordes de las imágenes eran irregulares y con manchas o degradaciones, físicas (raimientos diversos) y químicas (obscuraciones por degradación) o simplemente desbordamientos parciales de las imágenes sobre los teóricos bordes (como ya pasaba en los negativos del fondo SACE del ICC ya tratados en 1.1)



En esas circunstancias el recorte **automático** de los negativos se hacía muy difícil y se optó por una solución del tipo “aguas arriba”; en nuestro caso se diseñaron unas escuadras de **metacrilato transparente de color verde intenso** de forma que el operador ponía el original entre dichos soportes de manera que el borde quedase tapado por el reborde de la escuadra,



Imagen sin positivar

de esa forma el algoritmo de recorte era simple y 100% fiable – nótese que el operador puede controlar por transparencia la situación del borde a recortar –



Imagen positivada

De otra parte, a pesar de que los vidrios, así dispuestos, descansaban sobre la mesa de luz, al tener ésta su superficie tamizada, dispersaba la luz y no se producían los indeseados anillos de Newton.

Fondo Casas

Constituido por unos 10.660 vidrios negativos de distintos formatos, cuya especial dificultad estribaba en que algunos, en uno de sus bordes, había una pestaña oscura con una hendidura en la que podía haber información o no. Debiéndose entonces recortar preservando dicha información o no.



Resultando muy complicado el decidir cuándo hay o no hay información en esa zona.

Ya es bien complicado delimitar la zona, teniendo en cuenta que los bordes acostumbran a estar raídos, con manchas de distintos grises, como para decidir cuando el contenido de la zona es “información relevante” o cuándo son manchas o cuándo es suficientemente homogénea como para decidir que no hay información.

Finalmente se optó por un sistema semiasistido, con umbrales de seguridad que, de no rebasarse, dirigían la decisión a un controlador humano.



2.- DIFERENCIAS CONCEPTUALES ENTRE LA FOTOGRAFÍA CLÁSICA Y LA DIGITALIZACIÓN MASIVA DE DOCUMENTOS

La fotografía es, probablemente, una mirada subjetiva de la realidad, y que intenta reflejar una particular interpretación de la misma. Por eso una fotografía puede mover al espectador a tomar una concepción distinta de la escena, de la que tomaría en verla directamente.

Esa mirada el fotógrafo la resuelve en dos tiempos:

- en la propia toma y
- en el “revelado” posterior, en el cual hace emerger su “mensaje”. Revelado que acostumbra a realizar desde el formato RAW directo (en crudo) de la cámara o si no lo hará, a través de Photoshop, desde el formato ya revelado (TIFF, JPEG, etc.).

Por lo tanto, la fotografía es diferente para cada fotógrafo y para cada objeto fotografiado.

Sin embargo, en la digitalización masiva cada fotografía se hace de forma idéntica para cada objeto fotografiado dentro de un conjunto homogéneo.

Podríamos decir que es notarial, puesto que no es permisible que el digitalizador retoque la imagen de forma subjetiva para resaltar o atenuar aspectos que él considere.

Ese aspecto “notarial” merece algunas matizaciones

Es relativamente frecuente que las imágenes para difusión deban “mejorarse” en función de diversos objetivos, como por ejemplo que pesen menos a la hora de transferirlas por la red pero manteniendo su legibilidad o comprensión por parte del usuario

De cualquier forma esos retoques se deberán de hacer de forma automatizada, puesto que el coste de ir retocando manualmente cada fotografía sería inasumible. Lo cual no quiere decir que se aplique lo mismo en cada imagen: el sistema deberá ser suficientemente “inteligente” como para adaptar sus retoques y sus parametrizaciones para cada imagen según él “vea”.

Hay un tipo de soportes que sufren patologías muy fuertes debidas a la “edad” y a una deficiente conservación, de tal manera que entonces sí que se nos pide que las imágenes máster también estén retocadas - lo cual consideramos que es un error, todo sea dicho de paso -

Éste es el caso de algunos acetatos, vidrios, negativos, etc., que pueden estar muy velados unos, muy oscurecidos otros y también algunos que están velados en unas zonas y oscurecidos en otras.

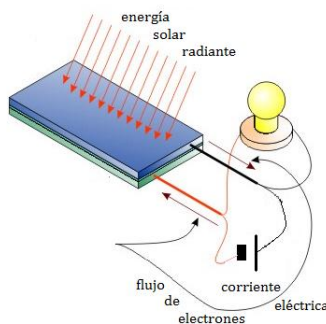
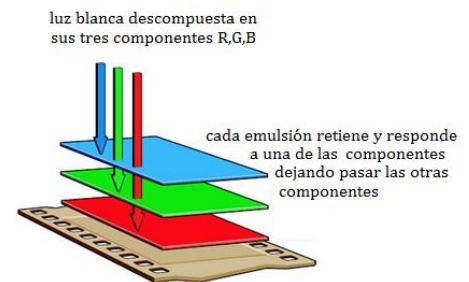
Pero muchas de las obsesiones de los fotógrafos siguen rigiendo en el ámbito de la digitalización:

- Análisis y fijación de las condiciones de entorno:
 - Por ejemplo la uniformidad de la iluminación
- Condiciones metodológicas previas a las tomas:
 - Calibración / corrección del color
 - Calibraciones / correcciones geométricas y geométrico-cromáticas
 - ...
- Evaluación de la calidad en la fase de pruebas.
Por lo que respecta especialmente a las copias de difusión, dado que las evaluaciones del apartado anterior han debido de fijar la calidad de las imágenes máster.
- Control continuado de las operaciones
- ...

Pero para que sus conocimientos se incardinan correctamente en los trabajos de digitalización, deberían conocer varios aspectos de ese tipo de trabajos masivos.

3.- LOS SENSORES FOVEON v.s. MOSAICS.

Detrás del conjunto de lentes de una cámara analógica clásica hay (había) un negativo sensible a la luz; de una sola capa de emulsión fotosensible, si se trataba de un carrete solo en grises, o si se trataba de fotografía en color, con tres capas de emulsión cada una sensible a una componente de color, una al rojo (red, R), otra al verde (Green, G) y la otra al azul (blue, B), tal como se muestra en la figura

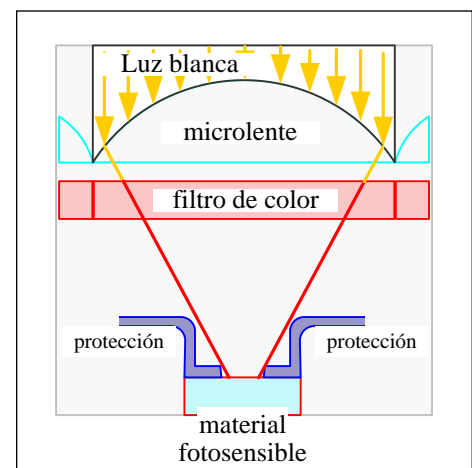


En las cámaras digitales, los sensores no son “granos de emulsión” sino que se trata de fotodiodos sensibles a la luz que cuando se someten a una radiación electromagnética como la luz emiten electrones que se recogen en una corriente eléctrica. O en términos más conocidos, podríamos decir que se trata de algo similar a una célula fotoeléctrica.

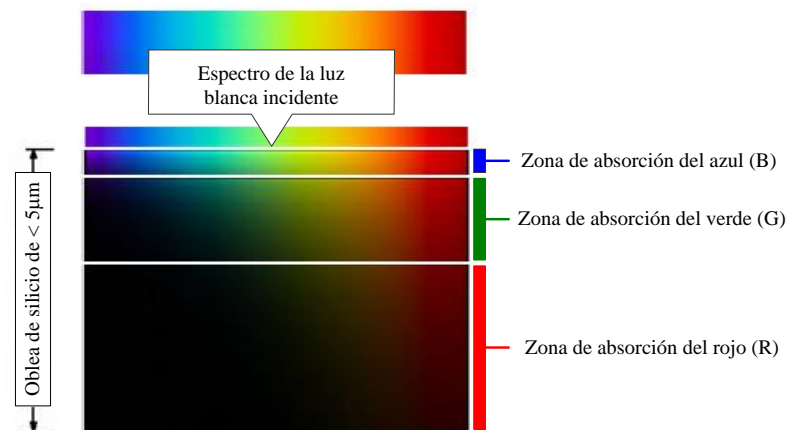
Una visión más cercana a la realidad de las células fotosensibles, que llamaremos **fotositos**, de las cámaras sería la siguiente

Delante del material fotosensible se coloca una microlente que dirige la luz hacia él y si se quiere que el sensor solamente responda a una de las tres componentes, se coloca un filtro intermedio que retenga el resto de componentes no deseadas.

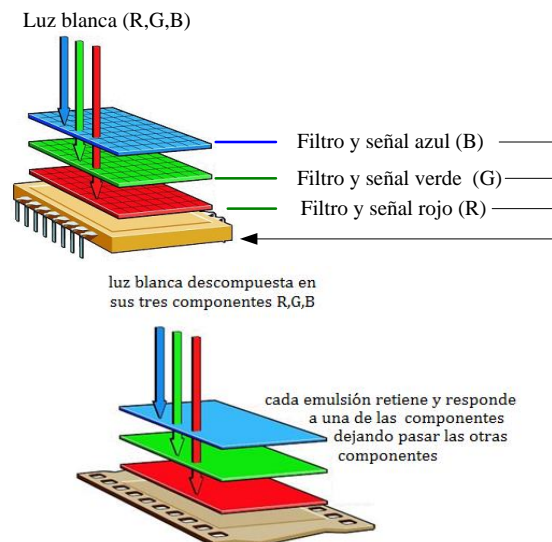
En ambos casos la corriente eléctrica que se genere será función de la intensidad de la luz recibida y será trasladada de alguna forma a formato digital, normalmente con una longitud de 12 o 14 bits y por lo tanto abarcará una gama de valores que irá del 0 al 4.096 (12 bits por fotosito) o del 0 al 16.384 (14 bits por fotosito)



Pero algunos materiales, como las obleas de silicio, son filtros progresivos que van absorbiendo diferentes longitudes de onda a medida que la luz va atravesando mayores espesores, hasta convertirse en totalmente opacas. La figura adjunta ilustra dicha absorción progresiva

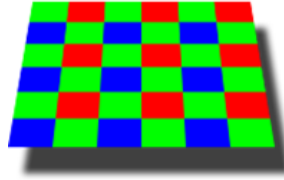


Los sensores Foveon están compuestos por una matriz de fotositos que aprovechan esta propiedad. Es como si debajo de cada zona de absorción hubiese un material fotosensible de manera que cada célula dará un valor para cada una de las tres componentes primarias de la luz. Así, la matriz Foveon recuerda la disposición de capas de emulsión de las clásicas películas para foto analógica en color, tal como se ilustra a continuación



Así, en los sensores Foveon, cada fotosito da las tres señales R,G,B que corresponden, en su conjunto, a la intensidad de la luz incidente. A cada uno de esos conjuntos (R,G,B) le llamaremos **píxel** - acrónimo del inglés "picture element" - y el conjunto de píxeles ordenados en la misma forma matricial será lo que llamaremos **imagen digital**.

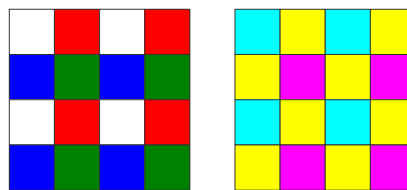
En cambio, los sensores de tipo mosaico siguen siendo matriciales; pero cada fotosito solamente da idea de la luz de una sola de las tres componentes de la luz que le llega, bien sea la R, la G o la B, para lo cual delante de cada célula, además de la lente le interpone un filtro de color, bien sea rojo, verde o azul. El sensor de tipo mosaico más usado es el Bayer RGBG, que dispone los filtros R,G y B, siguiendo el siguiente mosaico



Hay el doble de sensores dedicados al verde que al rojo o al azul en imitación del ojo humano, cuya distribución de conos “verdes” es, más o menos, el doble que la de los conos “rojos” o de los “azules”.

Ahora, en realidad, cada píxel (cada fotosito) solamente refleja una de las componentes de la luz incidente.

Hay otros tipos de mosaico menos usados, como por ejemplo el Bayer RGBW o el CYYM, entre muchos otros:

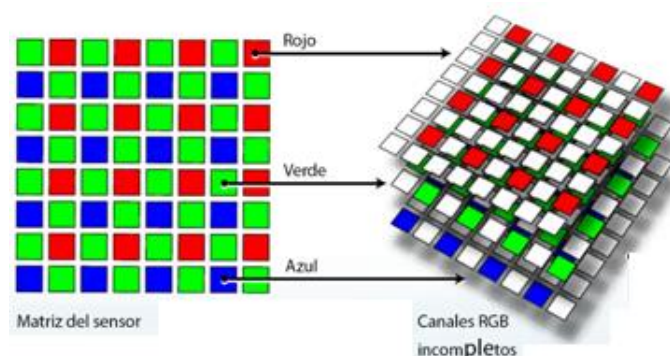


Bayer RGBW

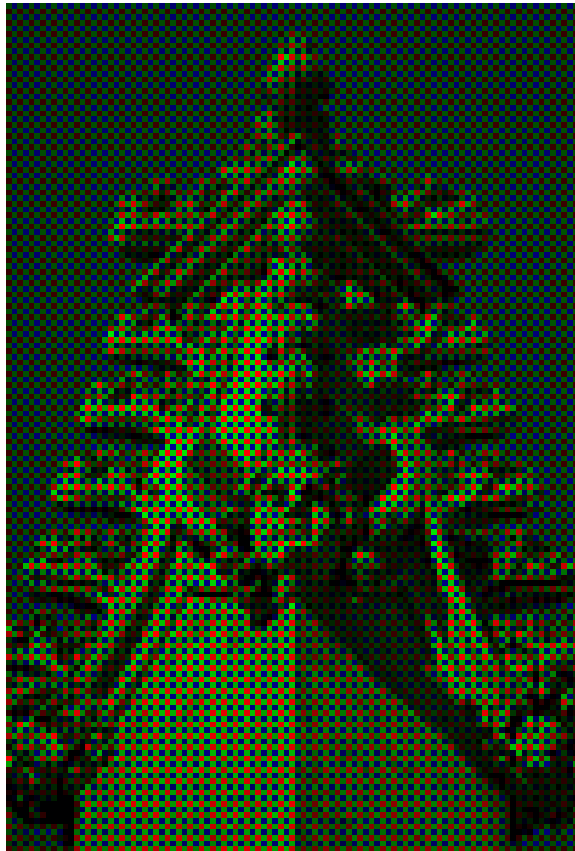
CYYM

El primero incorporando células que serán sensible al conjunto “blanco”, que por lo tanto será redundante o utilizando los primarios sustractivos cian, amarillo y magenta (duplicándose el amarillo).

En cualquier caso, las distintas señales de los fotositos de un sensor en mosaico deben combinarse con las de sus vecinos de forma que podamos interpolar las componentes que faltan en cada célula. Es como si tuviésemos tres imágenes; pero cada una con “agujeros” de valor desconocido



Si la diésemos por buena, la imagen podría tener un aspecto como el siguiente

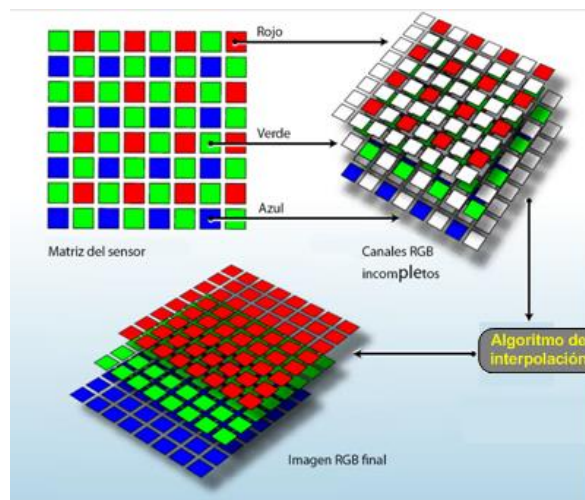


Hemos de implementar pues un algoritmo que interpole los agujeros de cada capa, de manera que finalmente asignemos a cada fotositos tres valores RGB completos para conseguir un píxel completo (R,G,B). Algoritmo que se conoce como “**demosaicing**” o informalmente como “**debayer**”.

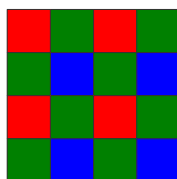
4.- EL PROCESO DE “DEMOSAICING”. IMPLICACIONES EN EL CONCEPTO DE RESOLUCIÓN

4.1.- El proceso de “demosaicing”

Decíamos que el conjunto de algoritmos que pasan de las tres “imágenes virtuales” de los tres canales de un mosaico a una sola imagen recibe el nombre de “**demosaicing**”



Así por ejemplo, un algoritmo básico de interpolación para un Bayer RGBG podría ser el siguiente:



R_{11}	R_{12}	R_{13}	
R_{21}	R_{22}	R_{23}	
R_{31}	R_{32}	R_{33}	

G_{11}	G_{12}	G_{13}	G_{14}
G_{21}	G_{22}	G_{23}	G_{24}
G_{31}	G_{32}	G_{33}	G_{34}
G_{41}	G_{42}	G_{43}	G_{44}

	B_{22}	B_{23}	B_{24}
	B_{32}	B_{33}	B_{34}
	B_{42}	B_{43}	B_{44}

$$\begin{aligned} R_{12} &= (R_{11} + R_{13}) / 2 \\ R_{21} &= (R_{11} + R_{31}) / 2 \\ R_{22} &= (R_{11} + R_{13} + R_{31} + R_{33}) / 4 \\ R_{23} &= (R_{13} + R_{33}) / 2 \\ R_{32} &= (R_{31} + R_{33}) / 2 \end{aligned}$$

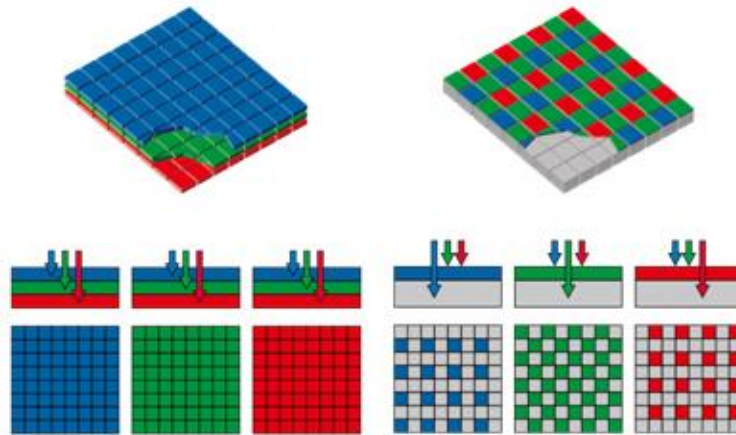
$$G_{22} = (G_{12} + G_{21} + G_{23} + G_{32}) / 4$$

$$\begin{aligned} B_{23} &= (B_{22} + B_{24}) / 2 \\ B_{32} &= (B_{22} + B_{42}) / 2 \\ B_{33} &= (B_{22} + B_{24} + B_{42} + B_{44}) / 4 \\ B_{34} &= (B_{24} + B_{44}) / 2 \\ B_{43} &= (B_{42} + B_{44}) / 2 \end{aligned}$$

Aunque en general se usan algoritmos un tanto más sofisticados, con entornos más amplios a tener en cuenta (en el algoritmo de nuestro ejemplo hemos usado entornos de 3x3) y se complementan con filtros para minimizar artefactos y mejorar las transiciones entre zonas de un color a otro.

4.2.- Implicaciones en el concepto de resolución

Si decimos que la resolución de una cámara es el número de píxeles de la imagen digital que genera, tenemos que:



En los Foveon, cada fotosito recoge la luz de una parte de la escena y la convierte en tres señales, R,G y B, que la describen totalmente, o sea que la convierte en un píxel; mientras que en los de tipo mosaico, esta misma parte de la escena llegaría a un fotosito y éste solamente sería capaz de darnos la descripción de una de las tres componentes, según de qué color sea la microlente que tenga delante de su película sensible.

Podríamos decir que un fotosito Foveon vale por tres de uno en mosaico. O al revés, que un fotosito de un sensor de tipo mosaico vale 3 veces menos que uno Foveon. Como que los de mosaico, si tienen Q millones de fotositos dicen que tienen Q millones de píxeles completos, pero si los Foveon siguiesen esa forma de ver su resolución deberían decir que tienen si tienen Q millones de sus fotositos dicen, con razón, que tienen $Q \times 3$ millones de píxeles – a su tecnología la llama “Foveon X3” precisamente por esa razón – pero hemos de tener claro que la imagen digital resultante de esa cámara Sigma no tendrá de hecho $Q \times 3$ millones de píxeles, sino tan solo Q. Por el contrario, en las de mosaico debermos tener en cuenta que sus Q megapíxeles provienen de una interpolación de colores.

Aun y así, las ventajas de los Foveon parecen claras, y efectivamente consiguen una **mejor transición de colores**, pero tienen dos inconvenientes importantes:

- 1.- Las dificultades de fabricación que redundan en un precio más caro
- 2.- Solamente la marca de cámaras Sigma monta Foveon, lo cual reduce las opciones en prestaciones, mientras que el resto de cámaras montan Bayers y por lo tanto las hay de todos tipos, prestaciones y precios.

Actualmente se detectan varios movimientos tendentes a emular las matrices de tipo Foveon por parte de varios constructores.

De todas formas, mientras tanto, nosotros solamente trataremos sensores de tipo Bayer RGBG

5.- LA RESOLUCIÓN. UN EJEMPLO DE DISONANCIAS ENTRE FOTÓGRAFOS, IMPRESORES Y DIGITALIZADORES

La resolución es un concepto ubicuo que se aplica a una gran cantidad de cosas y con muy diferentes sentidos.

- Una **lente** tiene una cierta resolución, y por lo tanto un conjunto de lentes también.
- Una **película** tiene su propia resolución como resultado de las resoluciones de cada una de sus capas.
- Los **sensores** de una cámara digital también tienen cada uno de ellos una determinada resolución debido a su naturaleza (CCD, CMOS, ...), a su tamaño individual, a la separación entre cada uno de ellos dentro de la matriz de sensores, etc.
- Una **impresora** también tiene su resolución dependiendo del número de puntos que puede imprimir en una pulgada
- Un **monitor** tiene su resolución en función del número de puntos luminosos que es capaz de activar
- ...

Todas estas diferentes resoluciones pueden englobarse dentro de lo que se denomina la **MTF** (Modulation Transfer Function) de cada uno de los elementos citados o de cada una de sus diferentes geometrías. En definitiva, se trata de medir lo que se “**pierde**” en cada una de las acciones: en el paso de cada lente, en la química de cada capa de la película, en la relación señal/ruido de cada sensor, en su capacidad de respuesta a cada intensidad y a cada frecuencia de la luz incidente, a las pérdidas de geometría por aberraciones de distorsión, de colorimetría por aberraciones cromáticas, etc.

Sin ninguna duda todas esas resoluciones tienen una importancia capital en la calidad de la imagen resultante, pero los constructores de las cámaras no dan información de cada uno de esos aspectos y ya no digamos de la información que dan los constructores de escáneres que incorporan esas cámaras, con lo cual no podemos centrarnos en estos parámetros para valorar uno u otro escáner.

Es cierto que de esas resoluciones podemos hacer una medida indirecta, como mínimo en su conjunto, por ejemplo podemos ver cuál es el límite en la separación de dos líneas paralelas para que se pierda la capacidad de reproducir esa separación, o cuál es el diámetro de un punto para que se pierda del todo o solo parcialmente, etc.; pero entonces toparemos con otras resoluciones a tener en cuenta para poder determinar cuantitativamente esas resoluciones que queremos medir: la resolución de las pantallas y/o de las impresoras que usamos para generar los patrones de test (impresoras) y luego para determinar los resultados (pantallas u otra vez impresoras). Lo mismo pasará si queremos analizar cuantitativamente el posible viñeteo y demás aberraciones de cada uno de los elementos del sistema.

Todo ello a la vez constituye un lío gordiano para el cual, ni los fotógrafos ni los digitalizadores, tienen en general espada que empuñar, puesto que un análisis riguroso de cada uno de esos aspectos es realmente muy costoso en tiempo y en medios materiales. En general nos conformamos con una serie de pruebas sobre originales estándar y analizamos los resultados más o menos cuantitativamente hasta concluir qué cámara y qué objetivo nos dan un mejor resultado para cada caso. Y eso tan solo lo pueden hacer aquellos digitalizadores que pueden variar las cámaras y los objetivos de sus escáneres, lo cual es poco frecuente puesto que la mayoría solo disponen de escáneres que montan una determinada cámara, mejor o peor, y un solo objetivo, mejor o peor, sin posibilidad de cambiar nada.

Pero volvamos a los problemas con la resolución.

5.1.- La resolución según los digitalizadores

Los digitalizadores trabajan con cámaras fijas y con documentos que están (idealmente) a una distancia también fija de la cámara. La escena que captan se reduce al documento en cuestión con una pequeño margen a su alrededor que desdramatice el posicionado de los distintos documentos a digitalizar en el área de visión de la cámara. En esas condiciones si la escena **captada** tiene $E_x \cdot E_y$ pulgadas y el sensor tiene $F_x \cdot F_y$ fotositos y por lo tanto nos acabará dando $F_x \cdot F_y$ píxeles, cumpliéndose obviamente que

$$\frac{E_x}{E_y} = \frac{F_x}{F_y} \left(\text{para lo cual es necesario que los sensores sean cuadrados} \right) \\ \text{y hoy en día lo son la práctica totalidad}$$

Entonces el digitalizador dice que está digitalizando los documentos a una resolución de

$$\rho = \frac{F_x}{E_x} = \frac{F_y}{E_y} \text{ píxeles por pulgada (ppp)}$$

Lo cual indica que está dedicando ρ píxeles para resolver en imagen digital una pulgada lineal de la escena, o sea del documento que es la parte de la escena que finalmente le interesa.

5.2.- La resolución según los fotógrafos

Los constructores de cámaras casi siempre usan la palabra resolución para indicar el número de millones de píxeles que retorna finalmente su cámara (aunque sea después de algoritmos de demosaicing y filtros diversos). Así por ejemplo la Nikon D810A tiene, **redondeando**, una resolución de 36 millones de píxeles efectivos (36 Mp) dado que su sensor está compuesto por $7360 \times 4912 = 36.152.320$ sensores y genera, de una u otra forma, 36 Mp.

Así los fotógrafos (algunos, algunas veces) creen que la resolución al digitalizar con una cámara de 36 millones de píxeles será de 36 millones de píxeles. Pero eso no dice nada de qué tan bien se estará digitalizando el documento, lo cual dependerá de la distancia entre la cámara y el documento. Si la aleja suficientemente del documento, la imagen que de él dará la cámara será ilegible, reducida a unos pocos píxeles.

Eso es así de obvio, pero en algunos concursos, el pliego de prescripciones “técnicas” manda que se digitalicen los documentos a una resolución de tantos megapíxeles, sin más, de forma que el digitalizador usará una cámara de esas características y la pondrá a la distancia que buenamente quiera de los documentos a digitalizar o, en el mejor de los casos, los libros de horas los digitalizará a una resolución enorme mientras que para los planos DinA0 la resolución será de todo punto de vista insuficiente.

5.3.- La resolución según los impresores

En el mundo de la reproducción, la calidad de esa reproducción depende, entre otros muchos factores, del número de “manchas de color”, también llamados “puntos”, “dots” o también píxeles, que la máquina reproductora puede imprimir por cada pulgada lineal del papel de impresión, que es a lo que ellos llaman resolución de la impresora. Así, hoy en día, cualquier impresora da, como poco, 600 píxeles por pulgada:

600 i (dots per inch), 600 ppp (puntos o píxeles por pulgada)

:: Aunque los equipos más profesionales alcanzan resoluciones mucho más altas (superiores a los 2000 ppp) ::

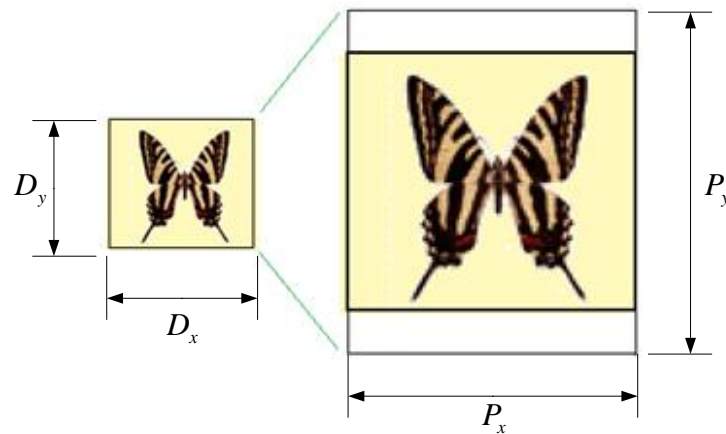
Desde ésta óptica, algunos de esos profesionales piden que la digitalización se haga, por ejemplo, a 300 ppp; pero lo que quieren en realidad es que la imagen del documento que se les dé, cuando la impriman en una impresora a una resolución de impresión de 300 ppp les ocupe en DinA3 (o un DinA4, o un 18x24, etc.).

Ese modo de definir la resolución impone al digitalizador la realización del siguiente cálculo:

Si se nos pide una “resolución” tal que al reproducir el documento en un papel de $P_x \cdot P_y$ pulgadas con una impresión de resolución de μ dpis, eso quiere decir que se deberán imprimir tantos píxeles como:

$$\begin{aligned} &P_x \cdot \mu \text{ píxeles en la dirección de las } x \\ \text{y} &P_y \cdot \mu \text{ píxeles en la dirección de las } y \end{aligned}$$

Pero eso solo será posible si las proporciones del documento son las mismas que las del papel a imprimir. O sea, que si nuestro documento tiene unas dimensiones de $D_x \cdot D_y$ pulgadas, debería cumplirse que $\frac{P_x}{P_y} = \frac{D_x}{D_y}$; pero eso casi nunca se cumple, de forma que en una de las dos dimensiones quedará papel en blanco.



A menos que deformemos la imagen a imprimir, lo cual es inadmisibile.

La situación es pues que tenemos un papel de dimensiones $P_x \cdot P_y$; un documento de dimensiones $D_x \cdot D_y$ y un sensor de $F_x \cdot F_y$ píxeles. Hemos de ver a qué resolución, en términos de píxeles por pulgadas del documento, hemos de realizar esa digitalización.

Anotemos:

$$\begin{aligned} P &= \max \{P_x, P_y\}; \quad p = \min \{P_x, P_y\} \\ D &= \max \{D_x, D_y\}; \quad d = \min \{D_x, D_y\} \\ F &= \max \{F_x, F_y\}; \quad f = \min \{F_x, F_y\} \end{aligned}$$

Al digitalizar el documento a una resolución de ρ ppp, la imagen resultante será de

$$\rho D \cdot \rho d \text{ píxeles}$$

y debería cumplirse que

$$\rho D = \mu P, \text{ y además que } \rho d = \mu p;$$

pero si las proporciones del papel de impresión y del documento original no son iguales, para asegurar que no haya desbordamiento en uno u otro lado, deberemos tomar

$$\rho = \mu \cdot \min \left\{ \frac{P}{D}, \frac{p}{d} \right\} \text{ ppp (píxeles por pulgada)}$$

6.- EL BAYER, EL RAW Y EL TIFF

El formato RAW es el formato directo (en crudo) de las señales que aporta cada uno de los fotositos del sensor Bayer, antes de cualquier demosaicing. Por lo tanto, tendremos tantas señales como fotositos tenga el sensor Bayer, y cada fotosito nos entregará una señal codificada **bien en 12 o bien en 14 bits**, correspondiente a una sola de las componentes de la luz que haya incidido en cada fotosito (en las matrices de tipo Bayer)

Si el sensor Bayer dispone de Q millones de fotositos, el conjunto de las Q millones de señales, de 12 o 14 bits, antes de cualquier proceso de demosaicing o filtros de ningún tipo, más algunas informaciones de cómo se realizó la toma (apertura, tiempo de apertura, etc.), es lo que constituye el formato RAW. Por eso se le llama “formato en crudo” (“raw” en inglés significa crudo en castellano).

Cada constructor, y a veces cada modelo de cámara, tiene su **propio** formato RAW

Fabricantes y extensiones de algunos ficheros raw

Canon	.cr2 .crw	Sony	.srf .sr2 .arw
Nikon	.nef .nrw	Pentax	.pef .ptx
Hasselblad	.3fr .fff	Samsung	.srw
Panasonic	.raw .rw2	Fuji	.raf
Olympus	.orf	Sigma	.x3f

En una primera aproximación, la estructura de esos distintos formatos RAW se basan en un formato TIFF en el que se ubica:

- la imagen tal cual sale del sensor con una profundidad de 12 o 14 bits comprimida sin pérdida de información
- un miniatura de la imagen procesada, ya no en formato RAW; generalmente JPEG
- metadatos de la cámara como son parámetros de exposición, balance de blancos, etc. en formato EXIF.
- Una imagen en formato JPEG de la imagen procesada para pre-visualización,
- perfiles de color
- etc. (según constructores)

De forma que para abrir y tratar un determinado RAW hace falta un programa específico.

Adobe ha definido un formato RAW (**.DNG**) con la intención de estandarizar el formato, cosa que de momento no ha conseguido. También existe una asociación llamada **OpenRaw** cuyo objetivo es la preservación de las imágenes, para ello defienden el formato RAW por ser el que tiene toda la información, pero la hermeticidad de los fabricantes, a la hora de divulgar estos formatos, complica su uso para conservación.

Otros procesos que se pasan son filtros de reducción de ruido, corrección del balance de blancos, contraste, saturación de color, enfoque, etc. Todos estos procesos de conversión es lo que se llama “**revelado**” de la imagen RAW en analogía (nostalgia) de los revelados químicos de las películas analógicas.

Otro aspecto en el que se dan disonancias entre los fotógrafos y los digitalizadores es en los “revelados” o sea, en el paso del formato primigenio RAW al formato manejable TIFF. Las diferencias derivan fundamentalmente de las existentes entre un “revelado” manual, imagen a imagen, con una carga de subjetividad importante (sin que ello comporte carga peyorativa de ningún tipo), frente a un revelado “dinámico” aplicable a todos los originales.

Es verdad que, en algunos casos muy extremos, jugar con informaciones de 12 o 14 bpp puede extraer informaciones situadas en las franjas extremas que al truncar a los 8 bpp quedan anuladas; pero esos extremos pueden ser interesantes para tratamientos forenses, pero difícilmente serán de uso en temas documentales.

En general lo que pueda hacerse con el RAW puede hacerse con el TIF, siempre que el paso de RAW a TIFF se haya hecho correctamente – si al pasar de RAW a TIF le decimos al sistema que hemos usado una iluminación de una temperatura de color determinada y resulta que era de otra, el paso de RAW a TIF tomará ese parámetro y luego será difícil la reconversión desde el TIF. Peor será incluso si hemos realizado las fotografías con la detección del blanco en automático y luego queremos deshacer el desastre. Etc.

En definitiva, la digitalización masiva no puede realizar un “revelado”, o sea, una corrección tonal, imagen por imagen, sin encarecer en demasía el trabajo.

Lo que sí puede y debe hacer son correcciones tonales automáticas que corrijan en lo necesario ciertas patologías que puedan darse en los originales y que pueden impedir la comprensión del documento

6.1.- Algunos ejemplos de retoques dinámicos automatizados

Algunos ejemplos de retoques dinámicos automatizados

Retoques tonales sobre la gama de grises

Hay muchos originales que son grises o que fueron grises en su día: negativos antiguos, vidrios, etc. y que con el paso del tiempo presentan patologías muy agudas:

- .- están velados
- .- están ennegrecidos
- .- unas zonas están veladas mientras que otras están ennegrecidas
- .- ...

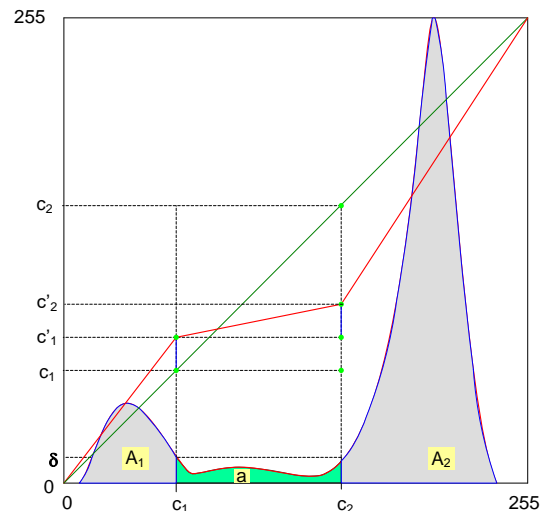
Los retoques suelen ser de dos tipos:

- .- retoques sobre la LUT en base a un análisis del histograma de grises
- .- retoques sobre zonas determinadas de la imagen

Los primeros pueden ser automatizados mediante sistemas dinámicos, los segundos solo en algunos casos. A continuación veremos el cómo y el porqué de unos y otros.

Retoques tonales de grises en base a retocar la LUT

Dado un histograma “reducido” de grises como el de la figura – reducido en el sentido que la frecuencia máxima la reduciremos hasta 255 y el resto de frecuencias quedarán reducidas de forma proporcional –



Al hacer la reducción de frecuencias podemos pintar la LUT en el mismo diagrama del histograma de los distintos niveles de gris, de forma que en principio la LUT será la diagonal que hemos pintado de verde y que corresponde con la LUT identidad.

Parece claro que la gama tonal de grises entre los tonos c_1 y c_2 está completamente desaprovechada puesto que con dichos tonos solamente se están representando un número de píxeles igual al área “a” que el histograma deja por debajo de él entre esos dos tonos (el área pintada de verde). Área que está claro que es mucho más pequeña que el resto de áreas cubiertas por el histograma A_1 y A_2 .

(– Está claro que el umbral δ que define el valle a “aprovechar” es un parámetro a determinar dependiendo de los objetivos –)

La idea es montar una LUT, como la dibujada en rojo, de forma que si antes para representar los píxeles del área “a” necesitábamos una gama tonal de c_1 a c_2 , ahora, con la nueva LUT, solamente dedicaremos una gama de c'_1 a c'_2 ; mientras que las gamas tonales que antes usábamos para representar los píxeles de las áreas A_1 y A_2 eran, respectivamente, de 0 a c_1 para A_1 y de c_2 a 255 para A_2 , ahora quedará ampliadas a las siguientes: de 0 a c'_1 para A_1 y de c'_2 a 255 para A_2 .

Obviamente deberemos determinar el escalón de c_1 hasta c'_1 : $(c'_1 - c_1)$ en función de a , A_1 y A_2 pero teniendo en cuenta que

$$a + A_1 + A_2 = W = \text{nº total de píxeles de la imagen}$$

podemos decir que el escalón $(c'_1 - c_1)$ será

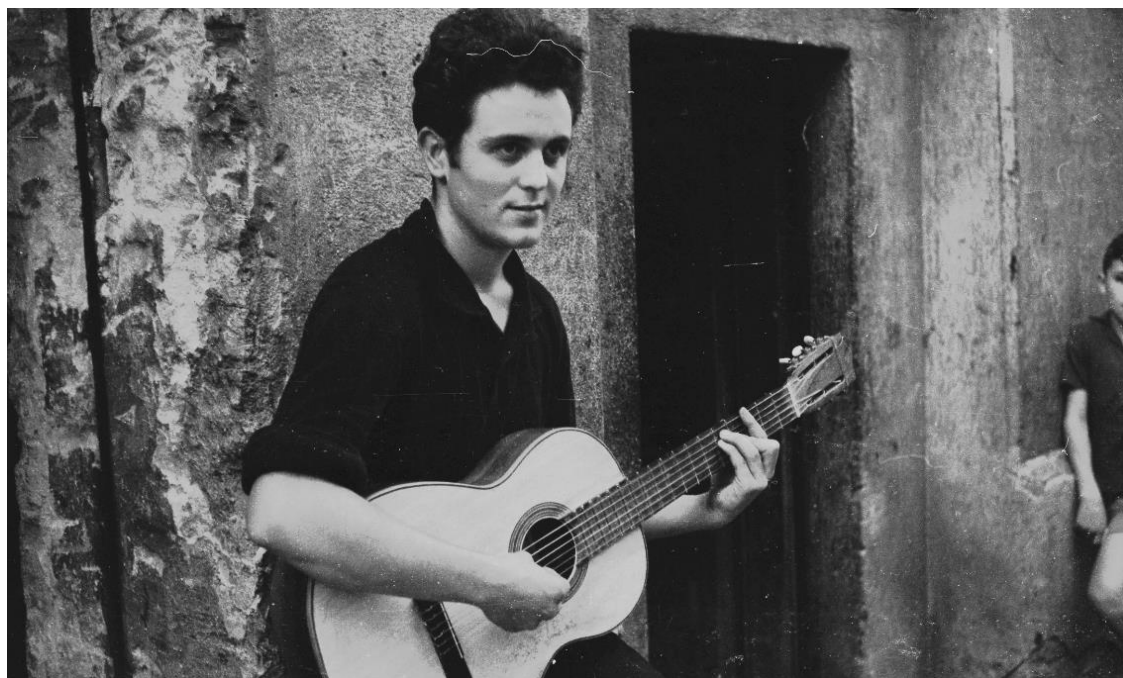
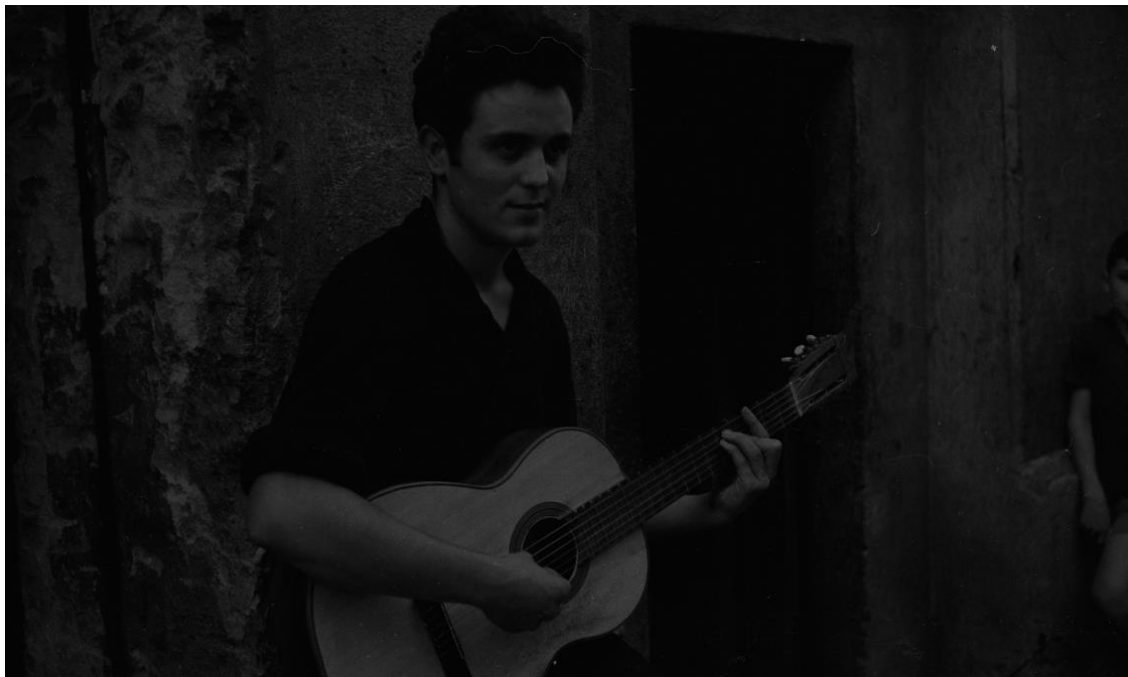
$$(c'_1 - c_1) = F(A_1, A_2, W)$$

Quedando la función F a determinar según las características del trabajo y de las patologías presentes.

Aquí, para simplificar la explicación, hemos “rectificado” la LUT mediante una línea quebrada, mientras que lo más adecuado es montar una spline y además la rectificación de ha limitado a un solo valle “desaprovechado”; pero ese proceso podría irse repitiendo, de izquierda a derecha del histograma, hasta que no quedase ningún valle por debajo del umbral δ .

Veamos un par de ejemplos de resultados





No siempre el histograma tiene la forma bimodal que nos ha servido para la presentación del sistema, de manera que el valle principal se encuentra a uno de los lados de la gama tonal (como es el caso de los ejemplos mostrados), en cuyo caso se realiza algo parecido a un estiramiento global del mismo.

Hay otras sofisticaciones a realizar posteriormente, como es la mezcla de las dos imágenes: la original y la procesada para usar las mejores zonas de una u otra, lo que podemos llamar el “mapeado zonal de tonos”.

De todas formas, está claro que **un operador** experto en Photoshop, o similares, puede conseguir **resultados todavía mejores**; pero, entre cargar la imagen y tantear hasta encontrar el mejor resultado y luego guardarlo, ello le puede costar unos cuantos minutos (pongamos 12 para simplificar los cálculos), que multiplicado por unos cientos de miles o millones de imágenes a procesar, le podría tener “entretenido” durante un buen “rato” toda varias vidas ($(10^6 \cdot 12 \text{ min})/60 = 200.000 \text{ h} = 25.000 \text{ días de } 8 \text{ h}$)

Binarizaciones – finalistas o intermedias –

La binarización de una imagen transforma a ésta en una imagen de tan solo blancos y negros puros, sin grises intermedios. Ésta binarización puede ser:

- **finalista**: cuando se quiere obtener como producto final esa imagen en blanco y negro para que sea más ligera, cosa bastante común en el ámbito documental.
- **Intermedia**: cuando algún proceso de mejora quiere apoyarse en el conocimiento de los “píxeles negros”

Aquí, en aras a la brevedad, solamente trataremos el primer caso, pero en cualquiera de ellos se tiene el problema de decidir **cuándo decidiremos que un píxel es blanco y cuando que es negro**.

■ Por umbral

El sistema más burdo, y el más frecuente, es decir que a partir de un umbral, por ejemplo a partir del gris del nivel 128, los píxeles que superen ese umbral los tomaremos como blancos y aquellos que no superen este valor los daremos como negros.

Los resultados acostumbran a ser horribles.

■ Según el histograma

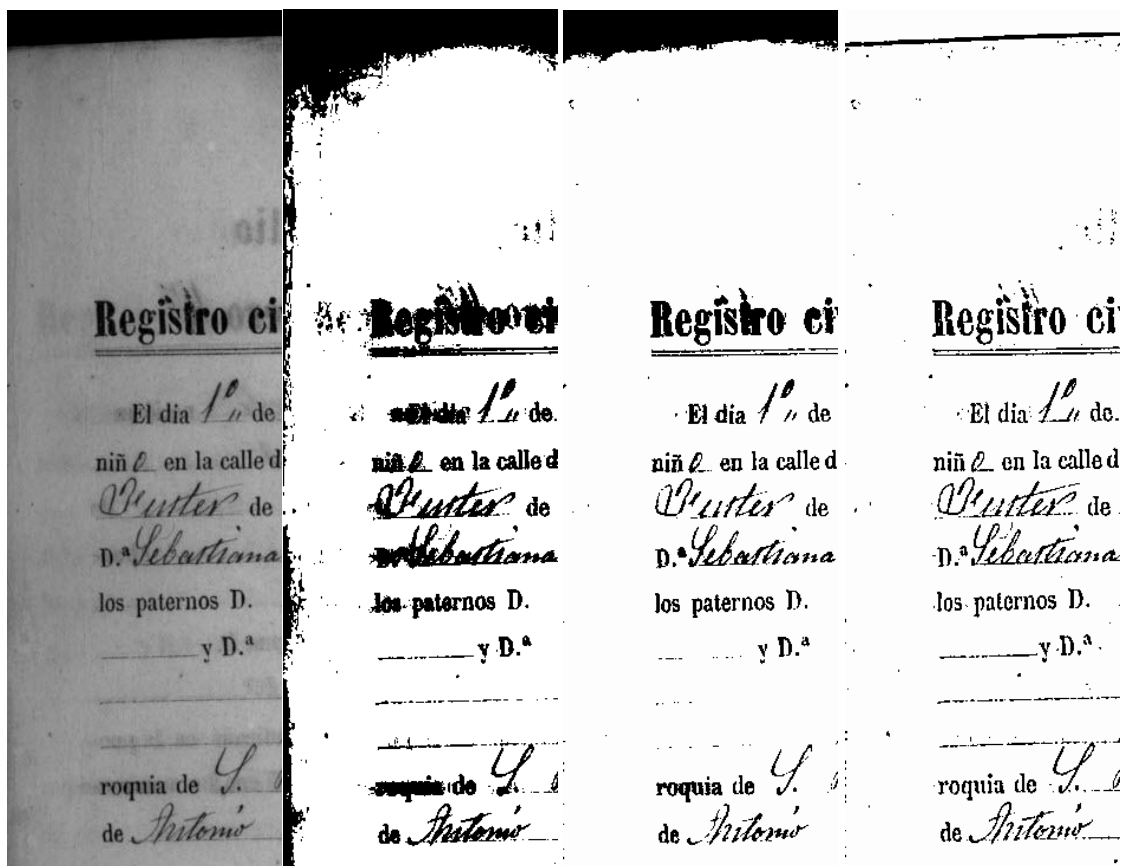
Un sistema más inteligente es acomodarse a la forma del histograma y determinar el umbral en función de dicha forma.

Los resultados son mucho mejores, pero en general insuficientes

■ Adaptación global-local

Otros sistemas analizan el entorno de cada píxel y toman la decisión sobre él en función de su valor en relación al de sus vecinos; pero así no se tiene en cuenta el estado general de la imagen, por si está sobreexpuesta o subexpuesta.

Así pues, parece claro que deberemos tener en cuenta de una parte la situación global de la imagen y de la otra la situación local del entorno del píxel a decidir. Así esos sistemas de binarización se pueden adaptar de forma “inteligente” a la situación de cada imagen y conseguir binarizaciones con pérdidas mínimas de legibilidad en el caso de los documentos.

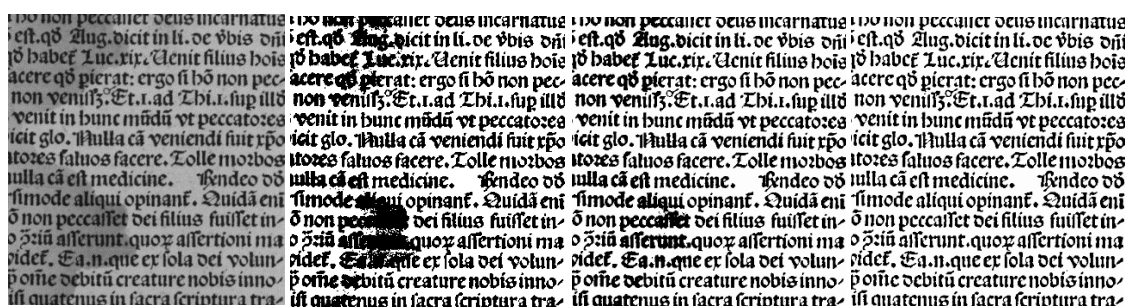


Original gris

Por umbral

Por histograma

Global-local



Original gris

Por umbral

Por histograma

Global-local

En definitiva,

sirvan esos ejemplos, tanto esas binarizaciones adaptativas, como las LUTs dinámicas, para ilustrar algunas diferencias entre el quehacer artesanal y artístico de los fotógrafos, y el trabajo masivo de los digitalizadores.

7.- LAS NUEVAS CÁMARAS DE 71 MP EN LA DIGITALIZACIÓN DE DOCUMENTOS

Adhoc siempre ha trabajado con cámaras matriciales, en vez de utilizar, como otros muchos, cámaras con sensor lineal que debe moverse para barrer la escena, entendiendo que ese barrido mecánico era muy lento y que su precisión podría ponerse en entredicho. Sin embargo, las cámaras lineales conseguían resoluciones más altas. Actualmente, la tendencia cada día está más decantada hacia las cámaras matriciales puesto que consiguen velocidades mucho más altas y resoluciones cada vez más competitivas.

Un caso paradigmático es las recientes cámaras “desnudas” de **71 Megapíxeles**.

Les llamamos “desnudas” porque se componen de una lente más un sensor matricial de 71 Mp y casi nada más.

- No disponen de pulsadores ni ningún otro elemento que permita operarlas.
- No tienen obturador ni partes móviles como espejos
- No disponen de dispositivo de autoenfoco
- ...

Éstas cámaras, en vez de tener en su interior, casi un ordenador completo como es el caso de las cámaras convencionales, **están manejadas completamente desde el ordenador:**

- El disparo se realiza mediante un comando del ordenador
- La apertura del diafragma se fija manualmente, así como el enfoque
- El tiempo de exposición se determina fijando el tiempo que transcurre entre una lectura del estado de “excitación” de los sensores y la lectura siguiente – entre una y otra se “resetea” el estado de los sensores –
- ...



Un ejemplo de cámara “desnuda” de 71 Mp puede ser la cámara la RMV71 de la casa Illunis.

La RMV71 se conecta al ordenador usando la interface Camera Link, ello requiere la instalación, en el ordenador, de una tarjeta capturadora de imagen (frame grabber), como por ejemplo la X64 Xcelera-CL de Teledyne-Dalsa. La conexión Camera Link permite el envío de comandos y la transferencia de la imagen. Los comandos nos sirven para cambiar parámetros, tales como,

tiempo de exposición, modo de captura (video o foto), formato de datos o velocidad de transmisión.

Una vez la imagen está en la tarjeta capturadora esta se encarga de hacer el proceso de demosaicing, transformando la imagen que nos llega en formato RAW a formato TIFF.



Sin embargo, Adhoc, finalmente, se ha decantado por la cámara X71 de Microbox, para sus tareas de digitalización de formatos mediano/grandes

<http://www.book2net.net/index.php/en/products/digitization-camera>

La cámara X71 de Microbox se conecta al ordenador a través del **bus USB3**, con lo que no requiere de hardware específico, como era el caso anterior y con el cual se consiguen velocidades de transferencia de 350 MB/s, con lo que conseguimos 3 imágenes por segundo. Siendo este mismo bus por el que viajan también los comandos de operación de la cámara.

El proceso de transformación de la imagen RAW a formato TIFF la realiza la CPU del ordenador.

La cámara ya incorpora algunas calibraciones, pero en cada proyecto hemos de hacer las siguientes:

- La uniformización de la iluminación
- La calibración de color, usando la carta de color X-Rite GretagMacbeth

Así como fijar los parámetros siguientes:

- El tiempo de exposición
- El enfoque – posee una herramienta de ayuda a la determinación del enfoque –
- Filtros de imagen

Su forma de construcción permite instalar diferentes sistemas de montaje de objetivos como F mount compatible Nikon, M mount compatible Leica, y Schneider Adapter. Permitiendo poner objetivos Shneider, Rodagon, Zeiss, Nikon, etc.

Especificaciones técnicas

Tipo de sensor	CMOS de 71Mp. Disparador electrónico
Tamaño del sensor	1000(h) x 7096(v) píxeles
Color	RGB Bayer color / escala de grises
Rango dinámico	63 dB
Velocidad de exposición	0,0001 – 2,0 segundos
Velocidad de captura	2,5 Imágenes/s en color
Bits por píxel	8, 12 o 16
Modo de captura	Disparo simple a ventana completa o fragmento y modo vídeo
Tipo de conexión	USB 3.0 (350MB/s)
Alimentación	12V =
Consumo	Pico 6W, reposo 1,8W
Lentes	Amplia gama de adaptadores para diversos formatos de montaje
Medidas	105mm x 125mm x 20,5mm
Peso	500g. sin adaptador de objetivo

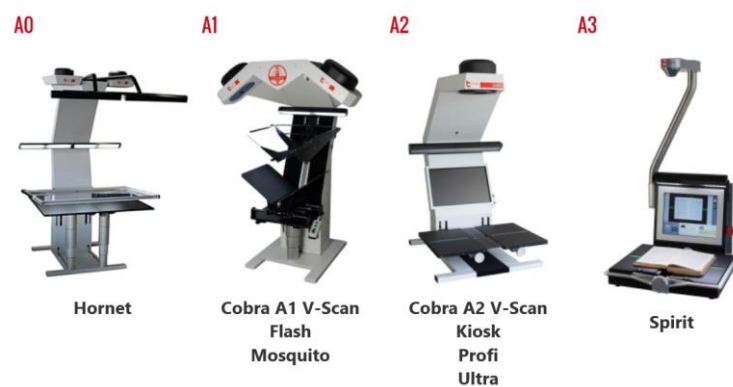
Estas cámaras están pensadas para aplicaciones específicas como son:

- **Digitalización de documentos**
- Reprografía
- Inspección industrial
- Vigilancia de grandes espacios
- Aplicaciones médicas y científicas
- Cartografía (fotogrametría aérea)

En nuestro caso las usamos, con un gran rendimiento, en la digitalización de documentos de todo tipo, cuando se requieren resoluciones altas y/o tamaños medianos/grandes.

<http://www.book2net.net/index.php/en/products/book-scanners>

La empresa Microbox GMBH de Frankfurt, en base a su cámara X71, ha desarrollado una serie de escáneres altamente competentes:



De la cual destacaremos, a título de ejemplo, el

Book2net KIOSK,

