



# Captura de imágenes digitales y preparación para su compresión

<b>Apellidos, nombre</b>	Oliver Gil, José Salvador (joliver@disca.upv.es)
<b>Departamento</b>	Informática de Sistemas y Computadores
<b>Centro</b>	Escola Tècnica Superior d'Enginyeria Informàtica

## 1 Resumen de las ideas clave

En este artículo se describe la formación de una imagen desde el punto de vista físico, y cómo se captura y digitaliza, dando lugar a un espacio de color RGB, que será recodificado en YCbCr con un submuestreo 4:2:0. Así lograremos una representación que requiera la mitad de memoria para su almacenamiento, sin apenas distorsión perceptible.

## 2 Introducción

El proceso de captura y digitalización de imágenes en color emplea habitualmente filtros con los que tan sólo se consigue información de uno de los tres colores primarios por cada elemento de la imagen (píxel), lo que después de un proceso de interpolación nos lleva de manera natural a obtener una representación de la imagen en un espacio de colores RGB.

El espacio de colores RGB no es generalmente el más adecuado a la hora de comprimir una imagen, ya que incluye una alta redundancia entre las tres componentes de color. Los planos rojo, verde y azul de una imagen presentan normalmente altas similitudes, lo que resulta poco adecuado cuando nuestro objetivo es comprimir la imagen. Por este motivo empleamos una representación en la que por un lado tenemos la componentes de luminancia de la imagen (intensidad de brillo) y por otro lado tenemos información adicional sobre el color, lo que además nos permitirá introducir una mayor reducción en la precisión de la información de colores, a la que el ojo humano es menos sensible. En realidad tampoco se transmite directamente intensidad de colores (es decir, de rojos, verdes y azules) sino niveles de crominancias, que es la diferencia entre cada uno de los planos colores y la componente de luminancia de la imagen. Con esto se consigue reducir también la redundancia existente dentro de una imagen entre su plano de luminancia y sus componentes de colores.

Todas estas acciones forman parte de una primera fase en cualquier escenario de compresión de imagen, en el que se elimina la redundancia que existe, a diversos niveles, en el espacio de colores RGB, que es el resultado natural del proceso habitual de captura de imágenes por filtrado de colores.

## 3 Objetivos

Una vez que el alumno lea con detenimiento este documento, será capaz de:

- Identificar los elementos clave del proceso completo de captura de imágenes digitales que posteriormente se emplean a la hora de comprimir una imagen.
- Realizar una conversión de espacios de colores, RGB a YUV y YCbCr con submuestreo 4:2:2 y 4:2:0.

## 4 Desarrollo

### 4.1 Definición física

Desde el punto de vista físico, las imágenes están compuestas por ondas electromagnéticas que tienen la particularidad de estar dentro de un rango que nos resulta visible a los humanos, como podemos observar en la Imagen 1. Sin embargo, no todos los colores se perciben igual, percibimos con mucha más intensidad los que se ubican en la parte central de este rango que los que están en los extremos próximos a infrarrojos o ultravioletas. Así, por ejemplo, dadas tres fuentes de luz de la misma intensidad, una verde, una roja y una azul, nuestro ojo percibe la luz verde con el doble de intensidad que la roja y seis veces más que la azul. ¿Por qué crees que sucede esto? Para resolver esta cuestión, ayúdate de la representación del espectro visible que se muestra en la Imagen 1 y fíjate en qué posición se ubica cada uno de los tres colores.

Efectivamente, la luz verde está ubicada en la parte central del espectro visible y las otras en zonas más próximas a los extremos no visibles. Recuerda este fenómeno porque te será útil más adelante.

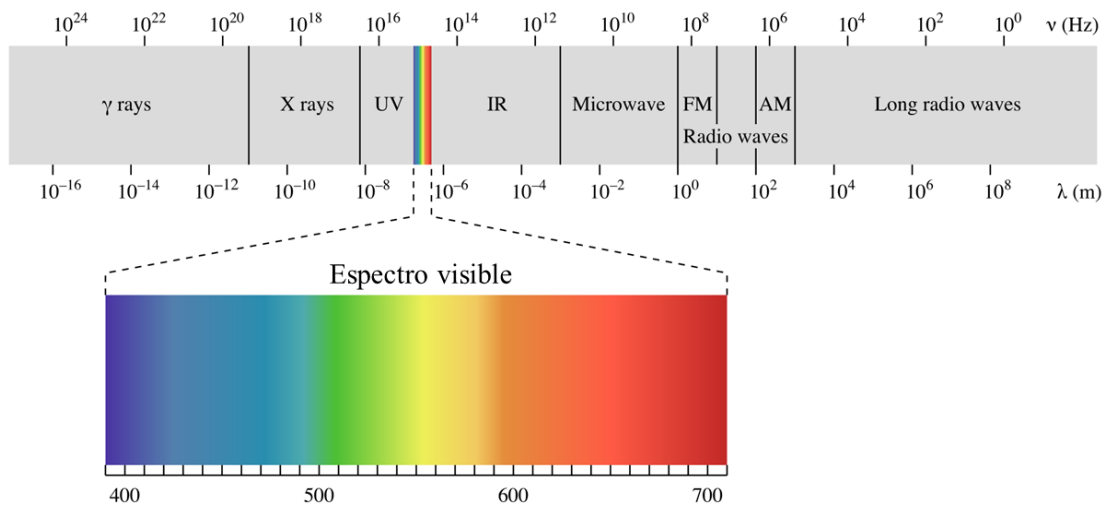


Imagen 1. Espectro visible

### 4.2 Captura de una imagen

Para conocer las características de las imágenes y así poder comprimirlas, resulta interesante observar cómo se capturan estas. Las cámaras digitales tienen sensores de tipo CCD o CMOS que se ubican en chips como este. La lente se encarga de situar la luz de la imagen a capturar en la ventana de este sensor, como muestra la Imagen 2.

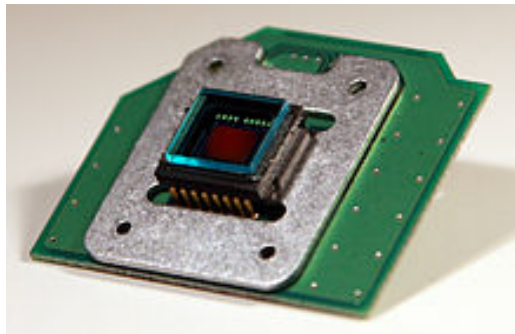


Imagen 2. Ventana de un sensor CMOS

Esta ventana tiene una serie de fotodiodos que convierten la luz en cargas eléctricas (es decir, fotones en electrones). En general, cada diodo del CCD da lugar a una muestra en la imagen capturada. A las muestras de las imágenes también se las denomina píxel, del inglés *picture element*.

El proceso de captura, así descrito, tiene un problema. Cada fotodiodo es capaz de captar intensidad de luz, pero no información de color, con lo que la imagen es similar a la de las antiguas televisiones en blanco y negro (formalmente llamada imagen en escala de grises o luminancia).

Para conseguir una imagen en color se usa un curioso mecanismo. Sobre cada uno de los fotodiodos se sitúa un filtro de color, de forma que algunos fotodiodos reciben información sobre rojos, otros sobre verdes y otros sobre azules. Puedes observar en la imagen del ejemplo que se capturan muchos más píxeles verdes que azules y rojos. ¿Sabes por qué? Así es, por la mayor contribución del color verde a nuestra percepción de intensidad de luz.

Prácticamente todos los colores se pueden definir a partir de sus componentes de intensidad roja, verde y azul. Así que, para que la información de color sea completa, necesitamos que cada píxel tenga un valor para cada una de estas tres componentes. Para esto, las otras dos componentes de color de un píxel se interpolan de los fotodiodos vecinos.

Obtenemos así las tres matrices que forman la imagen. Una con información sobre la componente roja, otra con la verde y otra con la azul. A cada una de estas matrices se las denomina planos rojo, verdes y azul. A este espacio de color lo llamamos RGB, de las siglas inglesas *red*, *green* y *blue*.

### 4.3 Digitalización de una imagen

El chip CCD nos proporciona valores analógicos, sin embargo, para su conveniente uso en sistemas digitales, es necesario un convertidor analógico-digital, que obtenga la representación digital de cada píxel a partir de la señal eléctrica generada por cada fotodiodo.

En este proceso de conversión, habitualmente se emplean 8 bits por cada componente, lo que sumando las tres componentes nos da como resultado 24 bits por cada píxel (dicho de otra manera, 24 bits/muestra). A esta codificación RGB se la denomina color verdadero (o *true color*) porque prácticamente define todos los colores que el ojo humano puede distinguir.

## 4.4 Espacios de colores

Como hemos visto, debido al uso de filtros de colores en la captura, las imágenes se obtienen directamente en el espacio RGB. Sin embargo, a veces otros espacios de colores podrían ser más convenientes. El espacio de color YUV representa la información de brillo de la imagen y la de color en planos separados.

Al plano con información de luminancia o brillo de la imagen se le denomina Y y su cálculo aproximado es un 30% de la componente roja, un 60% de la verde y un 10% de la azul, de la siguiente forma:

$$Y = 0.3R + 0.6G + 0.1B$$

¿por qué crees que se pondera de esta forma la contribución de cada uno de los tres colores a la luminancia? Si recuerdas, anteriormente dijimos que la componente verde se percibe con el doble de intensidad que la roja y seis veces más que la azul, por lo que son estas las proporciones que estamos empleando.

Junto con la luminancia debemos guardar también la información de color. Nos quedamos con la roja y la azul y descartamos la verde, al ser la más parecida a la información de luminancia y resultar redundante.

Las componentes azul y roja no se representan directamente mediante sus intensidades sino que se expresan como la diferencia de cada uno de estos colores respecto a la luminancia. A estas diferencias las llamamos crominancias. En concreto el plano de crominancia azul se le denomina U y el de crominancia roja se le denomina V. Es decir, se calculan así:

$$U = B - Y$$

$$V = R - Y$$

Fíjate que mediante esta diferencia se intenta reducir la redundancia existente entre los planos de color azul y rojo y el plano de luminancia, para facilitar el posterior proceso de compresión.

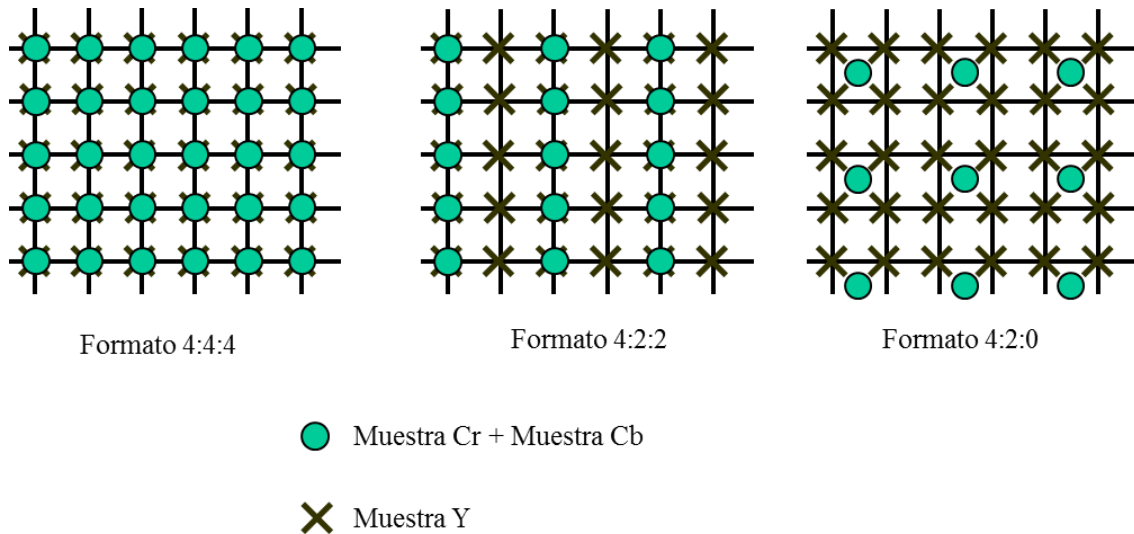
El espacio de colores YUV se emplea en señales analógicas, mientras que en las digitales usamos un espacio similar llamado YCbCr, que reescala las crominancias y les suma 128 para permitir una representación directa como byte sin signo. Su cálculo concreto es el siguiente:

$$C_b = (U/2) + 128$$

$$C_r = (V/1.6) + 128$$

## 4.5 Submuestreo de las crominancias

En el sistema visual humano, el ojo es muy sensible a errores introducidos en la información de luminancia, sin embargo, podemos perder precisión sobre la información de crominancia sin que se perciba demasiado. Una forma sencilla de realizar esto es mediante un submuestreo de los planos de crominancia (también llamado en inglés *subsampling* o *downsampling*).



*Imagen 3. Submuestreos habituales en el espacio de colores YCbCr*

Tal y como se observa en Imagen 3, si a una imagen no se le ha realizado ningún submuestreo decimos que está expresada en formato 4:4:4. En el formato 4:2:2 descartamos una de cada dos muestras en los planos de crominancia. En el formato 4:2:0, cada 2x2 muestras de crominancia nos quedamos con una (o sea, eliminamos la mitad de filas y mitad de columnas de cada plano de crominancia). Fíjate como, con este último formato, pasamos de tener tres planos completos a uno completo (el de luminancia) y dos cuartos de plano (cada plano de crominancia se ha reducido a la cuarta parte ya que tenemos la mitad de filas y la mitad de columnas). Por lo tanto, en el cómputo global hemos pasado de tres planos a uno y medio, con lo que reducimos a la mitad la cantidad de bytes necesarios para representar la imagen, sin percibir prácticamente deterioro respecto a la imagen original.

Ejemplos de estándares de compresión que emplean un submuestreo 4:2:0 son JPEG para imágenes, y H.261 y MPEG-1 para vídeo, mientras que MPEG-2, AVC y HEVC pueden usar tanto 4:2:0 como 4:2:2.

## 5 Cierre

Hemos visto el proceso completo de captura de una imagen y cómo se prepara la codificación para su posterior compresión haciendo uso de aspectos conocidos del sistema visual humano, como que nuestro ojo es más sensible a la información de luminancia que a la de crominancia y percibe con más intensidad las componentes verdes que las rojas y las azules.

## 6 Bibliografía

Fluckiger, F: "Understanding networked multimedia: applications and technology", Prentice Hall, 1995.



Oliver, J; Pérez-Malumbres, M: "Compresión de imagen y vídeo: fundamentos teóricos y aspectos prácticos", Ed. Universidad Politécnica de Valencia, 2001, pág. 13-18.