

Digitalización de imágenes y vídeo

[11.1] ¿Cómo estudiar este tema?

[11.2] Imágenes digitales

[11.3] Muestreo

[11.4] Cuantificación

[11.5] Resolución

[11.6] Vídeo analógico

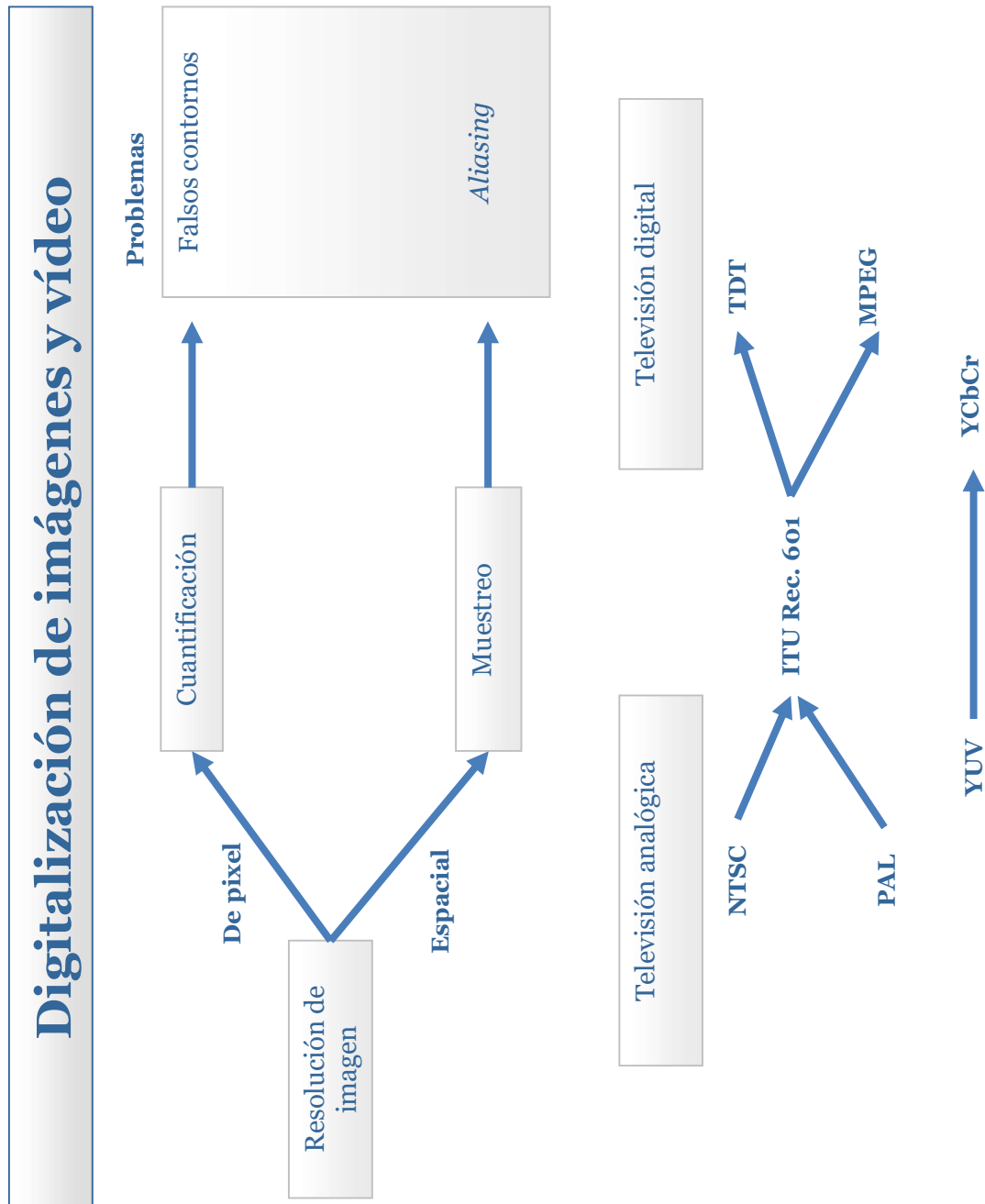
[11.7] Vídeo digital

[11.8] Vídeo en Octave

11

TEMA

Esquema



Ideas clave

11.1. ¿Cómo estudiar este tema?

Para estudiar este tema lee las **Ideas clave** que encontrarás a continuación.

En este tema veremos los principales parámetros que afectan a la digitalización de imágenes y vídeo. El objetivo es comprender cómo y cuándo estos parámetros afectan a la calidad de la imagen digitalizada. Al final del tema también se exponen los fundamentos del vídeo analógico y digital.

11.2. Imágenes digitales

La **digitalización de imágenes** permite representar imágenes analógicas del mundo real en un sistema digital. Una vez la imagen ha sido digitalizada resulta más sencillo su almacenamiento y procesamiento.

La digitalización de imágenes implica dos procesos:

- » **Muestreo:** donde se seleccionan un conjunto finito de puntos de la imagen analógica para su representación digital.
- » **Cuantificación:** donde se asigna un valor de amplitud discreta a cada punto, dentro de los infinitos valores que puede tomar la señal analógica.

El resultado de la digitalización es una **imagen**, es decir, un *array* o matriz rectangular de píxeles donde el valor de cada píxel corresponde a su intensidad (para imágenes monocromáticas) o a un vector de componentes de color (para imágenes en color).

Habitualmente se describe esta matriz de píxeles en **megapíxeles** (MP). Los megapíxeles son los millones de píxeles que tiene la imagen y se calcula como el producto del número de filas por el número de columnas que tiene el *array* de píxeles.

La figura 1 muestra los MP que ocupan diferentes resoluciones de imagen:

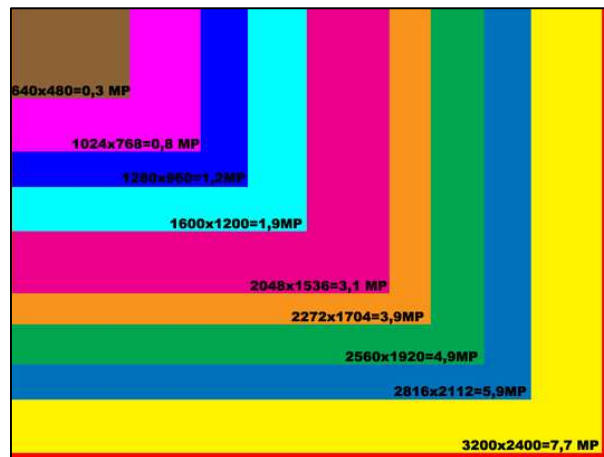


Figura 1: tonos sintéticos

Fuente: <http://pinstake.com/resoluciones/>

11.3. Muestreo

Una imagen se define como una matriz bidimensional en la que por convenio se fija el origen en la esquina superior izquierda, tal como muestra la figura 2.

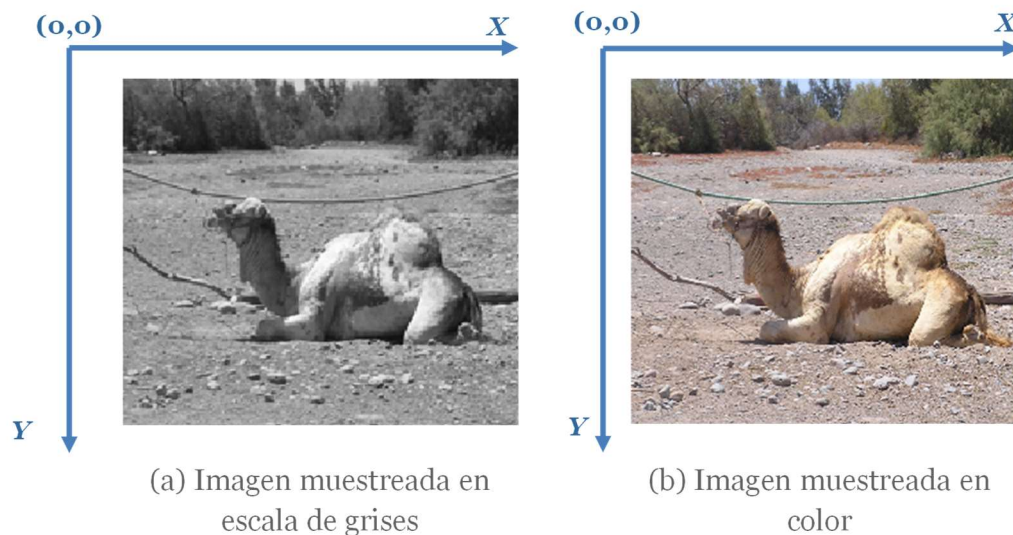


Figura 2: sistema de coordenadas de una imagen

Se pueden describir los valores muestreados como una función bidimensional $f(x,y)$ que nos da los valores de la imagen a intervalos discretos en el espacio. En el caso de imágenes de escala de grises, el valor de la función $f(x,y)$ representa la intensidad de luz detectada en ese punto. En el caso de imágenes en color, $f(x,y)$ representa un vector, con las tres componentes de color RGB.

Píxelado cuadrado

Un sistema de muestreo donde la densidad horizontal y vertical es la misma se dice que presenta **píxelado cuadrado**. En caso contrario se dice que tiene un **píxelado no cuadrado**. La figura 3 muestra un ejemplo de píxelado cuadrado y no cuadrado.

Observa que en la figura 3 (b) el ancho de los píxeles es mayor a su ancho:

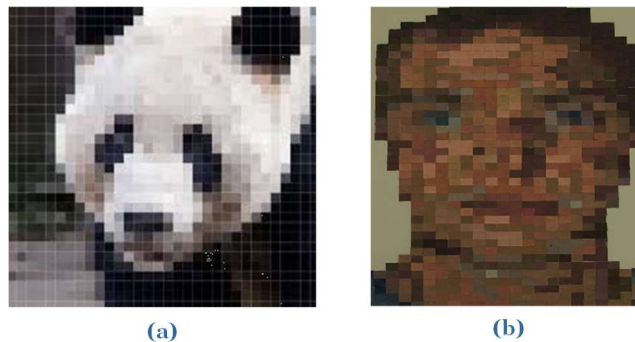


Figura 3: píxelado cuadrado y no cuadrado

Parámetros de muestreo

Los parámetros que deben ser tenidos en cuenta a la hora de muestrear una imagen son:

- » **Ratio de muestreo.** Indica el número de muestras que se toman por unidad física de ancho y alto de la imagen. Cuando mayor es el ratio de muestreo, mayor es la definición de la imagen. A este parámetro también se le llama **frecuencia de muestreo** pero el término «ratio» ayuda a recordar que se trata de un ratio entre las muestras tomadas y la longitud física de la imagen.
- » **Patrón de muestreo.** Indica la organización física de los puntos muestreados. El patrón de muestreo más utilizado es el rectangular, es decir, donde las muestras se distribuyen horizontal y verticalmente en filas y columnas, respectivamente. Más adelante describiremos otras organizaciones como la hexagonal y la polar-logarítmica.

Aliasing

El **criterio de Nyquist** dice que el ratio de muestreo debe ser de al menos dos veces la componente de máxima frecuencia. En caso contrario no habrá suficientes muestras para una reconstrucción adecuada de la señal. Si el muestreo se realiza a un ratio inferior, es decir, si hacemos un **inframuestreo** (*undersampling*) se produce ***aliasing***.

En 2D al *aliasing* se le llama **artefacto de Moiré** o **patrón de Moiré** y se caracteriza por la aparición de ondas o anillos extraños en las telas formadas originalmente por líneas paralelas.



Figura 4: patrones de Moiré

Fuente: <http://www.ishootshows.com/2012/04/09/understanding-moire-patterns-in-digital-photography/>

Este efecto se debe a la presencia de altas frecuencias en la dirección de las líneas final de la tela que han sido inframuestreadas al tomar la foto a una distancia insuficiente.

La figura 5 muestra cómo podemos generar el artefacto de Moire. La figura 5 (a) ha sido muestreada a una frecuencia suficiente como para no producir este artefacto. Por el contrario, la figura 5 (b) ha sido remuestreada a una frecuencia de muestreo por debajo del criterio de Nyquist, lo cual ha hecho que se produzca este artefacto en las zonas con mayor frecuencia espacial.

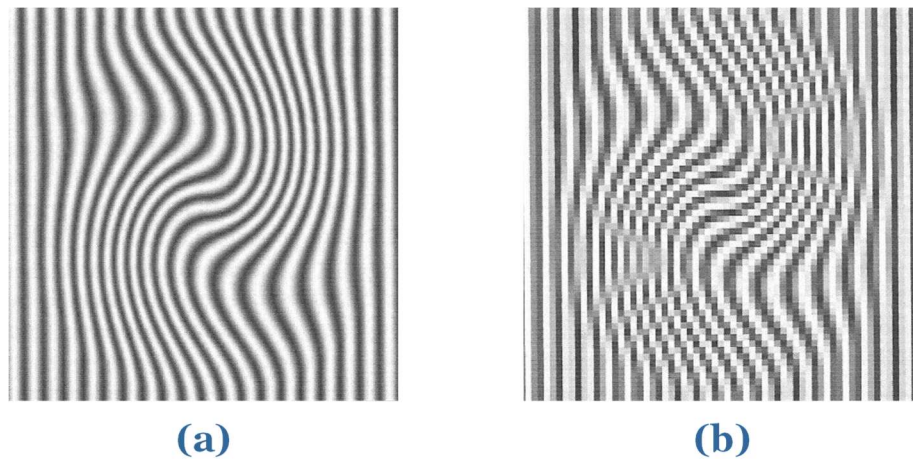


Figura 5: artefacto de Moiré simulado

Fuente: Efford, N (2000)

En el caso del vídeo también es posible producir un efecto de *aliasing* temporal conocido como el **efecto de las ruedas del vagón**, en el que los radios de la rueda parecen moverse al revés de la dirección de movimiento del tren. Esto se debe a que el número de muestras temporales de la rueda no es suficiente para la velocidad a la que se están moviendo los radios de la rueda.

11.4. Cuantificación

La cuantificación es el proceso de reemplazar el valor de una señal continua $f(x,y)$ por un conjunto discreto de **niveles de cuantificación**. La precisión con la que cuantificamos una señal continua $f(x,y)$ es mayor cuanto más niveles de cuantificación haya.

A los niveles de cuantificación muchas veces se les llaman **niveles de gris** porque representan diferentes intensidades en una imagen en escala de grises. En este caso, si hay n niveles de gris los niveles están representados por valores enteros $0, 1, 2, \dots, n-1$.

Por conveniencia n suele ser un entero potencia de 2, es decir:

$$n = 2^b$$

Donde b es el número de bits usados para cuantificar cada muestra. Habitualmente $b=8$, lo cual da lugar a $n=2^8=256$ niveles de gris.

Los sensores de algunas cámaras no son capaces de discernir más de 2^6 o 2^7 niveles de gris. Aun así, por conveniencia la muestra es convertida y representada a 2^8 niveles de gris, multiplicando por 2 o por 4 la muestra obtenida.

Los sensores de más calidad (como los usados en medicina) son capaces de discernir 2^{10} o 2^{12} niveles de gris. En este caso las muestras pueden ser subcuantificadas a 2^8 niveles de gris. También, para evitar perder información es muy habitual guardar las muestras con 2^{16} niveles de gris.

En el caso de las señales en color, cada nivel de cuantificación tiene asociados tres niveles, uno para cada canal RGB.

11.5. Resolución

La **resolución de una imagen** indica el nivel de detalle de una imagen.

Principalmente existen dos tipos de resoluciones: la resolución espacial y la resolución de píxel.

» La **resolución espacial** indica el tamaño físico de un píxel, es decir, la densidad de los píxeles de la escena. Cuanto más píxeles se use para representar un tamaño físico de la escena, mayor es la resolución espacial y mayor es la calidad de la imagen muestreada. La unidad de medida más habitual para la resolución espacial son los *dots per inch* (DPI), que indica cuántas muestras se toman en un segmento unidimensional de 1 pulgada (2.54 cm).

En ocasiones la unidad de medida DPI se usa para referirse a la resolución espacial de los dispositivos de salida (p.e. impresoras) y para medir la resolución espacial de los dispositivos de entrada (p.e. escáner) se usa otra unidad de medida llamada *pixels per inch* (PPI). Los PPI también se suelen usar en dispositivos de salida como los monitores, ya que estos dispositivos no generan puntos sino píxeles. Dado que estas unidades de medida representan el mismo tamaño, muchos autores las intercambian o usan indistintamente.

La figura 6 muestra el efecto de reducir la resolución espacial de una imagen con 256 niveles de grises. La forma en que se han generado estas imágenes se describe con más detalle en la lección magistral.

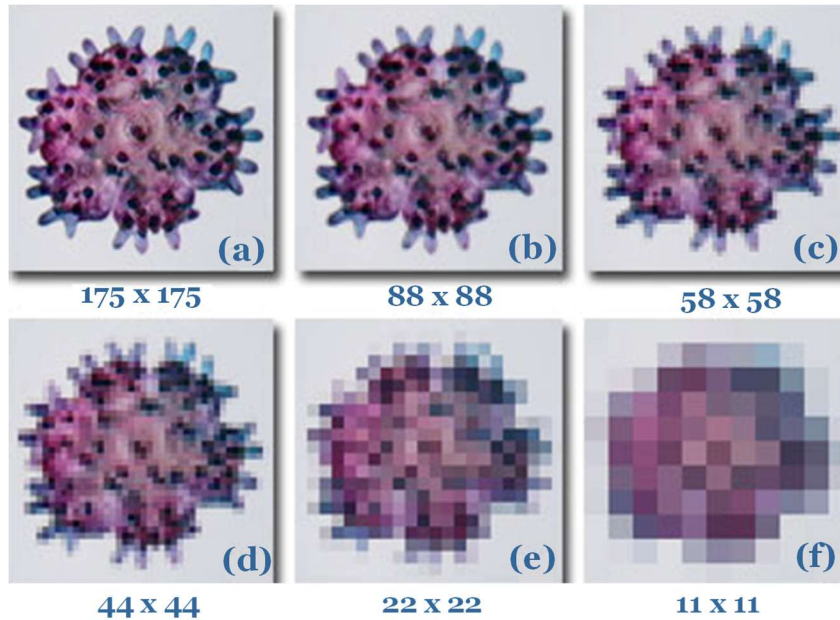


Figura 6: efecto de reducir la resolución espacial

Fuente: <http://www.olympusmicro.com/primer/digitalimaging/digitalimagebasics.html>



Figura 7: efecto de reducir la resolución espacial de grises

- » La **resolución de píxel** se refiere al número de niveles que se usa para cuantificar cada muestra de la imagen. En el caso de las imágenes de escala de grises la resolución de píxel también se llama **resolución de escala de grises** porque se refiere al número de niveles de grises utilizados en la cuantificación. En las imágenes en color la resolución de píxel se aplicaría a cada uno de los tres colores primarios.

La figura 7 muestra el efecto que se produce al reducir el número de niveles de gris de una imagen. Observar que la calidad de esta imagen es aceptable cuando tenemos 8 niveles de gris o más pero a partir de 4 niveles surgen lo que se llaman **falsos contornos** que hacen que se pierda la información de forma y que en consecuencia la resolución de píxel de la imagen se vuelva inaceptable.

11.6. Vídeo analógico

El proceso de escaneo y cuantificación

El **escaneo** es un mecanismo que mide una señal óptica de vídeo en cada instante de tiempo t y la guarda representada como una señal 1D eléctrica $f(t)$. La figura 8 muestra el proceso de escaneo de vídeo analógico.

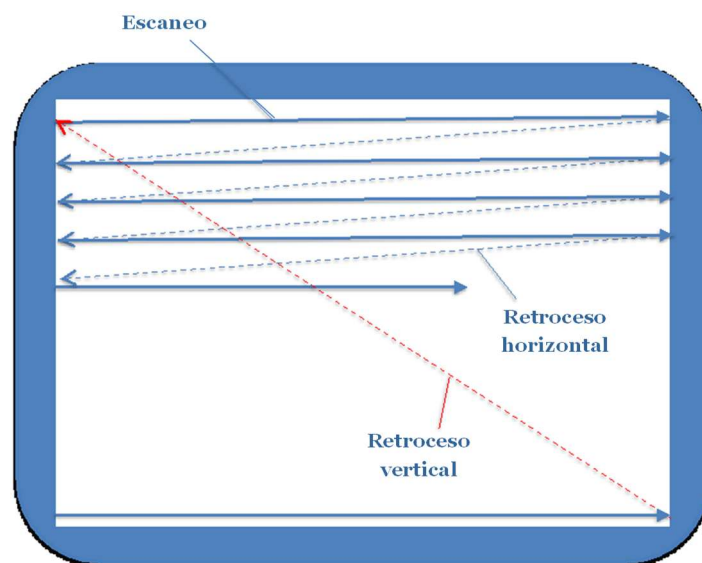


Figura 8: proceso de escaneo de vídeo

Las líneas gruesas corresponden al proceso de escaneo. Empiezan en la esquina superior izquierda y la línea de escaneo se mueve horizontalmente y en paralelo. Cuando el escaneo de una línea llega a la derecha, el proceso de escaneo se interrumpe y el lector se coloca rápidamente en la siguiente línea.

A este proceso se le llama **retroceso horizontal** (*horizontal retrace*): en la figura aparece marcado con líneas discontinuas azules. El retroceso está ligeramente inclinado hacia abajo con el fin de que el siguiente escaneo recorra una línea inferior.

Cuando se escanea la última línea se interrumpe el escaneo y se produce un **retroceso vertical** (*vertical retrace*), el cual aparece representado en la figura mediante una línea roja discontinua. Se llama **frame** al escaneo completo de una pantalla.

La secuencia de escaneo también ocurre cuando se reproduce el vídeo en un monitor analógico. La principal diferencia es que en vez de captar la señal $f(t)$ con un sensor, el monitor reproduce la señal $f(t)$ con un cátodo de luz.

Llamamos **escaneo progresivo** al escaneo de líneas una a una de abajo a arriba.

Aunque este es el proceso más lógico, las limitaciones técnicas que existían en los primeros sistemas de televisión analógica hicieron que se popularizara el escaneo entrelazado. En el **escaneo entrelazado** cada *frame* se escanea dos veces: primero las líneas impares y luego las pares. A cada pasada se le denomina **field** y contiene la mitad de líneas que el *frame*. En consecuencia, los *fields* se escanean en la mitad de tiempo que los *frames* y esto permitía reducir el efecto parpadeo que producía reproducir un vídeo con escaneo progresivo.

Formatos de vídeo analógico

Para que el ojo humano no perciba el parpadeo es necesario que los *frames* se muestren lo suficientemente rápido. Se llama **frecuencia de refresco** al número de *frames* que se muestran en un segundo. Se acostumbra a dar como *frames/seg* o como Hz.

Frecuencias de refresco típicas son:

- » **Cine.** 24 *frames/seg* entrelazado.
- » **TV.** 25-30 *frames/seg* entrelazado.
- » **Monitor de ordenador.** 75 *frames/seg* progresivo.

Se sabe que en la oscuridad no se percibe tanto el parpadeo, razón por la que en el cine se necesita una menor frecuencia de refresco (24 *frames/seg*). Observar que, al estar el cine entrelazado, la frecuencia de refresco es de 48 *fields/seg*.

Los formatos de vídeo analógico se acostumbra a clasificar con tres criterios:

- » Número de líneas horizontales escaneadas.
- » Frecuencia de refresco.
- » Indicación de si es entrelazado (2:1) o progresivo (1:1).

Los formatos de televisión analógica estándar SDTV (*Standard TV*) más usados han sido:

- » **NTSC**. Usado en EEUU y Japón. 525/60/2:1, que significa que hay 525 líneas a 30 Hz en entrelazado, es decir 60 *fields/seg*.
- » **PAL**. Usado en Europa. 625/50/2:1, que significa que hay 625 líneas a 25 Hz en entrelazado, es decir, 50 *fields/seg*

Posteriormente se diseñó una televisión de mayor resolución de pantalla a las que se llamó **HDTV** (*High Definition TV*). Este acrónimo no indica si la señal se transmite en analógico o en digital, sino el hecho de que la pantalla tiene mayor resolución.

Hubo varios intentos de HDTV analógica pero donde más aplicación ha tenido ha sido en la televisión digital. En la HDTV se utiliza una nomenclatura más compacta donde se omite la frecuencia de refresco y se concatena el número de líneas con la indicación de si es progresivo (*p*) o entrelazado *i*. Valores típicos son 720*p* y 1080*i*. Tanto en SDTV como en HDTV analógica, la frecuencia de refresco en EEUU es de 60Hz y en Europa de 50Hz.

Se llama **relación de aspecto** al ratio entre el ancho y alto de la pantalla. Valores típicos son:

- » 4:3 SDTV y monitores antiguos.
- » 16:9 HDTV y monitores nuevos.
- » 1.85:1 Cine.

Representación del color

La TV originariamente fue diseñada para transmisión en escala de grises. En consecuencia lo que se enviaba era la luminancia de los canales RGB. La luminancia se puede obtener de los canales RGB mediante la siguiente fórmula de ponderación:

$$Y = 0.299R + 0.587G + 0.114B$$

Cuando se desarrolló la TV en color se diseñó un sistema de representación del color compatible hacia atrás llamado YUV. En este sistema la señal de **luminancia** Y se envía por un canal y las señales de **cromanencia** $U=R-Y$, $V=B-Y$ se envían por otro canal adicional. Estas señales de color se calculan como:

$$\begin{aligned}U &= R - Y = 0.701R - 0.587G - 0.114B \\V &= B - Y = -0.299R - 0.587G + 0.889B\end{aligned}$$

Observar que la señal $G-Y$ no necesita ser transmitida, ya que puede obtenerse con la ecuación:

$$G = (Y - 0.3R - 0.11B) / 0.59$$

La razón por la que se decidió representar la diferencia de rojo $R-Y$ y la diferencia de azul $B-Y$ (y no la diferencia de color $G-Y$) es que el color verde G es el que más contribuye a Y , con lo que $G-Y$ es normalmente el número más pequeño y en consecuencia el más susceptible al ruido.

Curiosamente, cuando se tiene la misma capacidad de almacenamiento binario y el mismo nivel de ruido, la representación YUV es más robusta que la representación RGB.

Esto se debe a que el ojo es más sensible a las diferencias de luminancia que a las diferencias de color. Muchos estándares de codificación de vídeo representan con más precisión (o en mayor cantidad) las muestras de luminancia Y que las muestras de cromanencia U, V para conseguir una mejor calidad desde el punto de vista de percepción del ojo humano.

11.7. Vídeo digital

El **vídeo digital** es una representación del vídeo como una secuencia discreta digital.

Por contraposición, el vídeo analógico representa una señal continua. Los sensores CCD o CMOS son los encargados de capturar el vídeo digital como una secuencia de matrices bidimensionales de píxeles llamadas **frames**.

Las principales ventajas del vídeo digital frente al analógico son:

- » **Robustez.** A diferencia de la señal analógica, la señal digital no se degrada por la existencia de defectos o ruido en el medio de almacenamiento.
- » **Postprocesado.** Es más sencillo modificar y añadir efectos especiales al vídeo en formato digital.

Conversión de analógico a digital

En la sección anterior vimos que el vídeo analógico se representa como una señal 1D $f(t)$.

Por el contrario, el vídeo digital se representa como una señal 3D $f(x,y,t)$.

Si una señal de vídeo está capturada analógicamente necesitará ser digitalizada mediante un **conversor analógico-digital** (ADC). Este proceso implica básicamente tres pasos:

- » **Muestreo.** El ADC escanea a intervalos predefinidos de tiempo la señal analógica $f(t)$ y obtiene el valor de cada píxel (x,y) del *frame* digital, de la forma $f(x,y,t)$.
- » **Cuantificación.** El ADC asigna a cada señal muestreada un valor discreto dividiendo el rango de amplitud de la señal en intervalos.
- » **Codificación.** El ADC convierte los *frames* digitales en una representación más concisa mediante algún formato estándar de codificación de vídeo.

Parámetros del vídeo digital

Los *frames* del vídeo digital se caracterizan por:

- » El **framerate** (*fps*). El número de *frames* que se muestran por segundo.
- » El **número de líneas** (*h*). El número de líneas por *frame*.

- » El **número de columnas** (w). El número de muestras por línea.

Estas tres magnitudes se pueden invertir dando lugar a:

- » **Intervalo de *frame*** (Δ_t). El tiempo que permanece un frame en pantalla. Δ_t se puede calcular como $\Delta_t = 1/fps$.
- » **Alto de píxel** (Δ_h). Dada una pantalla de alto H , Δ_h se puede calcular como $\Delta_h = s_h/h$.
- » **Ancho de píxel** (Δ_w). Dada una pantalla de ancho W , Δ_w se puede calcular como $\Delta_w = s_w/w$.

Se llama **relación de aspecto del píxel** (PAR) al ratio entre el ancho y alto de cada píxel. Si un monitor tiene una relación de aspecto $AR=H/W$, la relación de aspecto de sus píxeles será:

$$PAR = AR \frac{w}{h}$$

La mayoría de los monitores de ordenador tienen $PAR=1$ pero existen estándares de TV como NTSC donde el $PAR=8/9$, o PAL donde el $PAR=16/15$.

Otro parámetro importante es la **profundidad de color** N_b , que indica el número de *bits* necesarios para representar un color. Un *frame* con 256 escalas de grises tiene una profundidad de color de $N_b=8$ *bits*, porque $2^8=256$. Si el *frame* tiene canales RGB, su profundidad de color será $N_b=24$ *bits*.

En vídeo, se llama **bitrate** (*bps*) al número de bits que se transmiten por unidad de tiempo. Si tenemos el *framerate* (*fps*), h líneas, w columnas y profundidad de color N_b podemos calcular el *bitrate* de una secuencia de vídeo descomprimido como:

$$bps = fps \cdot h \cdot w \cdot N_b$$

Representación del color

En vídeo digital el color se puede representar en RGB o en YUV. De nuevo, la ventaja de usar YUV es que las componentes de color se pueden representar con menor resolución espacial. Se llama **submuestreo del color** al proceso por el cual reducimos la resolución espacial de las componentes de color.

Cuando se trata de vídeo digital, al modelo de color YUV se le acostumbra a llamar YCbCr. Ambos corresponden al mismo modelo de color, con lo que nosotros usaremos la terminología YUV.

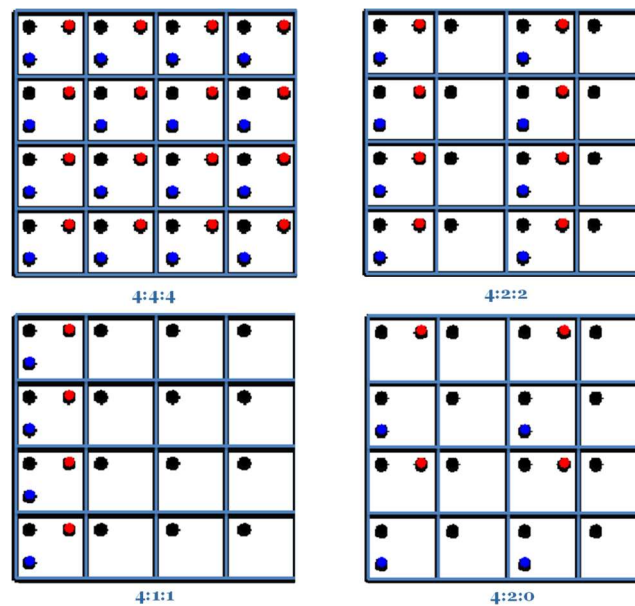


Figura 9: patrones de submuestreo de colores habituales

La figura 9 muestra los patrones de submuestreo más habituales. Los puntos negros corresponden a muestras de luminancia Y y los puntos rojos y azules a muestras de cromanencia U y V . En concreto serían:

- » **4:4:4.** No hay submuestreo, se toman tantas muestras de color como de luminancia.
- » **4:2:2.** Las componentes de color tienen la misma resolución vertical pero la mitad de resolución horizontal.
- » **4:1:1.** Las componentes de color tienen la misma resolución vertical pero la cuarta parte de resolución horizontal.
- » **4:2:0.** Las componentes de color tienen la mitad de resolución vertical y horizontal.
- » **Ejemplo.**

Si tenemos píxeles con profundidad de color $N_b=24$, calcular la profundidad de color para cada patrón de submuestreo cromático.

En el patrón 4:4:4 no submuestreamos, con lo que tenemos 8 muestras de color por cada 4 muestras de luminancia y $N_b=24$.

En el patrón 4:2:2 tenemos 4 muestras de color por cada 4 muestras de luminancia, con lo que tenemos $N_b=16$.

En el patrón 4:1:1 tenemos 8 muestras de color por cada 16 muestras de luminancia, con lo que tenemos $N_b=12$.

En el patrón 4:2:0 tenemos 2 muestras de color por cada 4 muestras de luminancia, con lo que tenemos $N_b=12$.

Esto quiere decir que en los patrones de submuestreo 4:1:1 y 4:2:0 se utiliza la mitad de bits para codificar un vídeo que en el patrón 4:4:4.

Formatos de vídeo digital

Existen multitud de formatos de vídeo digital. A continuación vamos a clasificarlos en función de su aplicación.

» Formatos para televisión SDTV.

Aunque la televisión SDTV es analógica, su producción y postproducción rápidamente empezó a hacerse en digital. El formato ITU Rec. 601 es el más utilizado para este fin. Fue diseñado para ser una versión digital compatible con los dos formatos de TV analógicos dominantes: NTSC y PAL. En consecuencia ITU Rec. 601 permite generar fácilmente vídeo para las parametrizaciones de ambos formatos analógicos.

Las parametrizaciones de estos formatos analógicos se resumen en la Tabla 1:

Parámetro	NTSC	PAL
<i>Frames</i> por segundo	30	25
<i>Fields</i> por segundo	60	50
Alto	720	720
Ancho	480	576

Tabla 1: parametrizaciones del formato ITU Rec.601

Asumiendo el patrón de muestro cromático 4:4:4 y observando los parámetros de la tabla podemos calcular el *bitrate* de NTSC y PAL. Observar que en ambos casos el *bitrate* resulta ser de 248.8Mb/seg:

$$NTSC_{bps} = fps \cdot h \cdot w \cdot N_b = 30 \cdot 720 \cdot 480 \cdot 24 = 248.8 \text{ Mb/seg}$$

$$PAL_{bps} = fps \cdot h \cdot w \cdot N_b = 25 \cdot 720 \cdot 576 \cdot 24 = 248.8 \text{ Mb/seg}$$

El submuestreo cromático 4:4:4 y 4:2:2 se utiliza habitualmente para producir y editar vídeo. Si usamos submuestreo 4:2:2 la profundidad de color baja a $N_b=16$ y el *bitrate* baja a 165.8Mb/seg. Sin embargo, para la distribución digital se suele utilizar el submuestreo cromático 4:2:0, lo cual significa que la profundidad de color baja a $N_b=12$ y el *bitrate* baja a 124.4 Mb/seg. El ejemplo más conocido de esta distribución digital de vídeo ITU Rec.601 es el DVD.

○ Ejemplo:

Si un DVD tiene capacidad para almacenar 4GB, ¿cuántos segundos de vídeo ITU Rec.601 puede almacenar?

4GB son $8 \text{ b/B} \cdot 4 \text{ GB} = 32 \text{ Gb} = 32768 \text{ Mb}$. Si dividimos la capacidad entre el *bitrate* obtenemos:

$$t = \frac{32768 \text{ Mb}}{124.4 \text{ Mb/seg}} = 263 \text{ seg} \approx 4 \text{ min}$$

Afortunadamente, el vídeo en DVD no se guarda directamente en formato ITU Rec-601 sino que se comprime con el formato MPEG-2, lo cual permite almacenar aproximadamente 120 minutos de vídeo.

» Tamaños de pantalla para aplicaciones de vídeo.

Para aplicaciones de codificación de vídeo tradicionalmente se han usado los tamaños de pantalla CIF, cuyos valores se resumen en la Tabla 2. El prefijo Q indica que QCIF incluye la cuarta parte de píxeles que el formato CIF. Observar, que al igual que el formato ITU Rec. 601, los formatos CIF son formatos con relación de aspecto 4:3. En temas posteriores se estudian formatos de pantalla más modernos.

Formato	Ancho	Alto
<i>Sub-QCIF</i>	128	96
QCIF	176	144
CIF	352	288
4 CIF	704	576

Tabla 2: tamaños de pantalla CIF

Debemos tener en cuenta que estos estándares CIF solo incluyen el tamaño de la pantalla pero no incluyen aspectos como el submuestreo cromático o el formato de codificación.

Sabemos que HDTV indica televisión de mayor resolución. En los formatos digitales se considera que una resolución es HDTV si tiene relación de aspecto 16:9 y 720p o más que corresponden con las tres últimas resoluciones de la **¡Error! No se encuentra el origen de la referencia..**

Señal de vídeo compuesta y de componentes

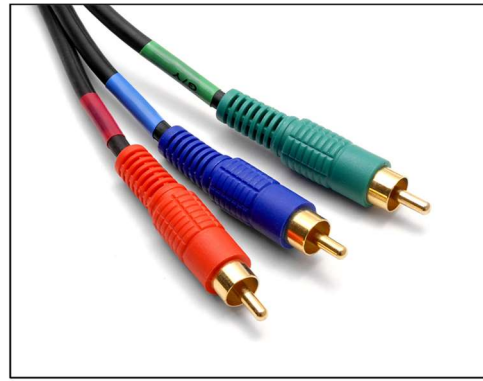
Se llama **señal de vídeo compuesta** a una señal de vídeo donde los tres canales de color (RGB y YUV) viajan multiplexados.

La ventana de la señal de vídeo compuesta es que solo se necesita un medio de transmisión. Por ejemplo, si se envía por el aire solo es necesario una frecuencia. Si se envía por cable solo se necesita un cable. La figura 10 (a) muestra un cable de vídeo compuesto.

Se llama **señal de vídeo de componentes** a una señal donde cada canal de color viaja por un medio de transmisión. La figura 10 (b) muestra un cable de vídeo de componentes. La ventaja del vídeo de componentes es que no se producen artefactos por interferencias entre los canales.



(a) Cable de vídeo compuesto



(b) Cable de vídeo de componentes

Figura 10: cable de vídeo compuesto y de componentes

Los sistemas de vídeo más tradicionales (como PAL y NTSC) usan señal compuesta, mientras que los sistemas más avanzados (como ITU Rec. 601 o al TDT) utilizan cable de componentes.

Habitualmente existen otros 2 cables para el audio izquierdo y derecho. Esto hace que las conexiones de vídeo compuesto tengan un total de 3 cables y las de vídeo de componentes un total de 5 cables.

Televisión Digital Terrestre

La **Televisión Digital Terrestre** (TDT) permite recibir señales digitales aprovechando las antenas de televisión analógicas convencionales, junto con un decodificador de TDT.

Su principal ventaja es un mejor aprovechamiento de espectro radioeléctrico. Este mejor aprovechamiento del canal se utiliza para mejorar la definición de la imagen y para aumentar el número de canales de televisión.

Los estándares de codificación han sido legislados por los gobiernos y existen multitud de ellos dependiendo de la zona geográfica o país. En Europa se implantó el formato DVB-T y en pocos años fue sustituido por el formato DVB-T2. Este segundo permite un 30% más de capacidad de canal debido al uso de formatos de codificación más eficientes.

La TDT emite **vídeo escalable**, es decir, vídeo que en función del tamaño de la pantalla del receptor utiliza parte o toda la información recibida por la antena.

Formato	Tamaño	Relación de aspecto
480i	720x480	4:3
480p	720x480	4:3
576i	720x576	45:36
576p	720x576	45:36
720p	1280x720	16:9
1080i	1920x1080	16:9
1080p	1920x1080	16:9

Tabla 3: formatos DVB-T2

La tabla 3 resume estos formatos. Observar que la relación de aspecto varía dependiendo del formato, siendo 16:9 la relación de aspecto de los de mayor resolución (HDTV). Los televisores de pantalla grande capaces de reproducir vídeo 1080p son conocidos como Full HD.

11.8. Vídeo en Octave

En esta sección vamos a ver cómo usar el *software* Octave para procesamiento de vídeo.

Este *software* usa matrices 3D para representar los *frames* de vídeo en escala de grises.

En el caso del vídeo en color las matrices son 4D porque la cuarta dimensión representa el color.

Alternativamente, si lo prefieres, puedes usar el *software* comercial MatLab.

Configuración del entorno

Para acceder a las librerías de vídeo en Octave debemos descargarnos el paquete «video».

Recuerda que para comprobar si está instalado debemos entrar en Octave y ejecutar el comando *pkg list*. Es decir:

```
$ octave
```

```
octave:1> pkg list
```

Package Name	Version	Installation directory
control *	2.8.3	/opt/local/share/octave/
general *	1.3.4	/opt/local/share/octave/
image *	2.2.2	/opt/local/share/octave/
signal *	1.3.2	/opt/local/share/octave
specfun	1.1.0	/opt/local/share/octave

Si no está instalado bájate (<http://octave.sourceforge.net/video/>) el fichero video-1.0.2.tar.gz y desde Octave puedes ejecutar el comando:

```
> pkg install video-1.0.2.tar.gz
```

Después debemos activar el paquete «video» con el comando:

```
> pkg load video
```

Leer ficheros de vídeo

Para consultar información sobre un fichero de vídeo podemos empezar usando la operación *aviinfo*:

```
> aviinfo('tortuga.mov')
```

```
Incorrect chunk size information in AVI file.
```

Observar que la función ha fallado porque el fichero no usa el formato AVI.

Podemos usar el comando *ffmpeg* para convertir el vídeo de ejemplo a este formato:

```
$ ffmpeg -i tortuga.mov -an -vcodec rawvideo -pix_fmt rgb24 -vf scale=640:320 tortuga.avi
```

Los parámetros de este comando piden que no haya audio (-an) y que el vídeo esté sin comprimir (-vcodec rawvideo -pix_fmt rgb24). Además en este caso lo hemos reescalado para hacer el *frame* más pequeño (-vf scale=640:320).

Ahora ya sí podemos consultar la información del fichero desde Octave. El comando `aviinfo` recibe el fichero y devuelve una estructura con campos describiendo el tamaño, formato, número de *frames*, ancho, alto, etc.:

```
> aviinfo('tortuga.avi')
ans =
    Filename: '/Users/flopez/tortuga.avi'
    FileSize: 80506184
    FileModDate: '07-Dec-2015 09:57:51'
    NumFrames: 586
    FramesPerSecond: 120
    Width: 640
    Height: 320
    ImageType: 'truecolor'
    VideoCompression: 'none'
    Quality: 4.2950e+07
    NumColormapEntries: 0
```

Al devolver una estructura podemos guardarla en una variable y luego acceder a sus campos individualmente. Por ejemplo:

```
> info = aviinfo('tortuga.avi');
> info.FramesPerSecond
ans =
    120
```

Ahora podemos cargar el fichero de vídeo con el comando:

```
> V = aviread('tortuga.avi');
```


Visualizar los *frames* leídos

En la variable *V* se habrá creado un vector fila con tantas columnas como *frames* tenga el vídeo:

```
> size(V)
ans =
    1 586
```

Cada elemento del vector es una estructura con dos campos:

- » *cdata* con la imagen del *frame*.
- » *colormap* una tabla de color o bien vacío si se usa color real RGB.

Luego si hacemos:

```
> V(1)
ans =
    cdata: [320x640x3 uint8]
    colormap: []
```

Vemos que el *frame* tiene 320 filas por 640 columnas y 3 colores RGB. En consecuencia el campo *colormap* está vacío.

Ahora podemos reproducir el primer *frame* usando *imshow* sobre *cdata*:

```
> imshow(V(1).cdata);
```

En este caso hemos accedido directamente a *cdata* porque sabemos que los *frames* usan color RGB. Si el *frame* estuviera representado en otro formato, debemos convertir el *frame* a imagen con el siguiente comando:

```
> I = frame2im(V(1));
```

El comando *frame2im* solo opera sobre un *frame*, con lo que si queremos convertir todos los *frames* del vídeo tenemos que hacer un bucle. Por ejemplo, si queremos aplicar un filtro de media espacial a todos los *frames* haríamos:

```
> for k=1:info.NumFrames
    I{k} = frame2im(V(k));
    I{k} = imfilter(I{k},fspecial('average', 4));
end
```

El uso de llaves en *I{k}* es para que *I* sea un *array* celda que es un *array* indexado que actúa como un diccionario. Si hubiéramos usado un *array* normal los elementos de *I(k)* deberían tener la misma dimensión que los elementos de *V(k)*.

Análogamente, si quisiéramos convertir una imagen a *frame* debemos hacer el proceso opuesto con el comando *im2frame*. Por ejemplo:

```
> for k=1:info.NumFrames
    V(k) = im2frame(I{k});
end
```

Escribir un fichero de vídeo

Una vez procesado el vídeo podemos escribirlo en un fichero con el comando *movie2avi*.

Por ejemplo:

```
> movie2avi(V,'tortuga_borrosa.avi');
```

Leer otros formatos

Aunque Octave no tiene esta opción, MatLab incluye y recomienda usar la operación *mmreader* que es más general y funciona con más tipos de ficheros. Esta función crea un reader multimedia que luego puede pasarse a la operación *read*:

```
> reader = mmreader('tortuga.mov');
> V = read(reader, [1 10]);
> size(V)
```

```
ans =  
    1080    1920     3    10
```

Esta operación nos carga una matriz 4D donde cada dimensión contiene:

- » Número de líneas.
- » Número de columnas.
- » Número de canales RGB.
- » Número de *frames*.

La variable *reader* es una estructura con información sobre el *frame*. Por ejemplo:

```
> reader.FrameRate  
ans =  
    26.6350
```

Para acceder a un determinado *frame* haríamos:

```
> imshow(V(:,:,,1));
```

También podemos reproducir todos los *frames* con el comando *implay*:

```
> implay(V,reader.FrameRate);
```

Lo + recomendado

Lecciones magistrales

Resoluciones espacial y de escala de grises

En esta lección magistral se estudia cómo usar Photoshop, GIMP e ImageMagick para modificar la resolución espacial y de píxeles de una imagen.



Accede a la lección magistral a través del aula virtual.

No dejes de leer...

Conceptos básicos de imagen digital

En este documento publicado por el Ministerio de Educación se describe la resolución y profundidad de color.

Diseño de Materiales Multimedia _Web 2.0

Accede al documento desde el aula virtual o a través de la siguiente dirección web:
<http://www.ite.educacion.es/formacion/materiales/107/cd/imagen/imagen01.html>

¿Cuál es la diferencia entre NTSC y PAL?

En este artículo se describen las diferencias entre estos formatos incluyendo zonas geográficas y evolución.



Accede al artículo desde el aula virtual o a través de la siguiente dirección web:

<https://curiosoando.com/cual-es-la-diferencia-entre-ntsc-y-pal>

No dejes de ver...

¿Qué es la resolución de una imagen?

En este vídeo de John Celis se explica qué es la resolución de una imagen, cómo se mide y cómo afecta a los píxeles.



Accede al vídeo desde el aula virtual o a través de la siguiente dirección web:

<https://www.youtube.com/watch?v=NP5R7IdiDJM>

+ Información

A fondo

La televisión digital

Cubero, M. (2009). La televisión digital: fundamentos y teorías. Barcelona: Marcombo.



Este libro profundiza en los estándares y formatos de vídeo y televisión.

Accede al artículo desde el aula virtual o a través de la siguiente dirección web:

<https://books.google.es/books?id=a8ZKGsUyWCoC>

Cine y televisión digital

Carrasco, J. (2010). *Cine y televisión digital. Manual técnico*. Barcelona: Ediciones Universidad de Barcelona.



En este libro se tratan los conceptos básicos del vídeo analógico y digital, la comprensión y la distribución.

Accede al artículo desde el aula virtual o a través de la siguiente dirección web:

<https://books.google.es/books?id=JEg4OfqRoagC>

Bibliografía

Akramullah, S. (2014). *Digital video concepts, methods and metrics: quality, compression, performance and power trade-off analysis*. Nueva York: Apress.

Cuevas, E. y García, A. (2010). *Procesamiento digital de imágenes con Matlab y Simulink*. Madrid: Ra-Ma.

Jack, K. (1993). *Video demystified: a handbook for the digital engineer*. Oxford: Newnes.

Martin, S. (2011). *Procesamiento digital de imágenes*. Madrid: EAE.

Weinard, D. (2015). *How video Works: from broadcast to the cloud*. Nueva York: Focal Press.

Test

1. Seleccionar un conjunto finito de puntos de una imagen analógica es:
 - A. Iterar.
 - B. Muestrear.
 - C. Submuestrear.
 - D. Operar.

2. ¿Qué tipo de función puede representar una imagen?
 - A. Lineal.
 - B. Logarítmica.
 - C. Exponencial.
 - D. Bidimensional.

3. El artefacto de Moiré se produce cuando:
 - A. La cámara no tiene una buena lente.
 - B. El muestreo sufre *aliasing*.
 - C. Hay líneas paralelas.
 - D. La imagen está curvada.

4. ¿Qué tipos de resoluciones son válidas para una imagen?
 - A. Espacial.
 - B. Temporal.
 - C. De píxel.
 - D. Espacial y de píxel.

5. Los falsos contornos aparecen:
 - A. Cuando inframuestreemos.
 - B. Cuando infracuantificamos.
 - C. Los dos anteriores.
 - D. Cuando se produce el artefacto de Moiré.

6. Se llama *frame*:
- A. A las líneas horizontales pares de un vídeo analógico.
 - B. A las líneas horizontales impares de un vídeo analógico.
 - C. A las líneas horizontales pares e impares de un vídeo analógico.
 - D. Ninguno de los anteriores.
7. ¿Qué *frame rate* tiene el cine?
- A. 24 Hz.
 - B. 25 Hz.
 - C. 26 Hz.
 - D. 30 Hz.
8. Si un vídeo tiene 720i significa que:
- A. Escaneamos 720 líneas horizontales.
 - B. Escaneamos 720 líneas verticales.
 - C. Escaneamos 360 líneas horizontales.
 - D. Escaneamos 360 líneas verticales.
9. ¿Cuál es la relación de aspecto de HDTV?
- A. 4:3.
 - B. 16:9.
 - C. 1.85:1.
 - D. 3:2.
10. ¿Qué patrón de submuestreo de color gasta menos ancho de banda?
- A. 4:2:2.
 - B. 4:1:1.
 - C. 4:2:0.
 - D. Varios de los anteriores.