

COPIA PRIVADA CON FINES DOCENTES

VISIÓN POR COMPUTADOR

ARTURO DE LA ESCALERA HUESO



3. Imágenes digitales

El concepto de imagen está asociado a una función bidimensional $f(x,y)$, cuya amplitud o valor será el grado de iluminación (intensidad de la luz) en el espacio de coordenadas (x,y) de la imagen para cada punto. El valor de esta función depende de la cantidad de luz que incide sobre la escena vista, así como de la parte que sea reflejada por los objetos que componen dicha escena. Estos componentes son llamados *iluminación* y *reflexión*, siendo descritos por $i(x,y)$ y $r(x,y)$ respectivamente. El producto de ambas funciones proporciona la función $f(x,y)$:

$$f(x,y) = i(x,y)r(x,y)$$

Siendo el intervalo de definición:

- $0 < i(x,y) < \infty$
- $0 < r(x,y) < 1$

por lo que $f(x,y)$ estará acotada:

$$0 < f(x,y) < \infty$$

La naturaleza de la iluminación viene determinada por la fuente de luz, mientras que la reflexión depende de las características del objeto en la escena.

3.1 Muestreo espacial y niveles de gris.

Una señal es una función que depende de alguna variable que tiene un significado físico. Dependiendo del número de variables puede ser unidimensional, o bidimensional, como el caso de las imágenes. Dependiendo de las dimensiones de la función pueden ser escalares (como las imágenes en blanco y negro), o vectoriales (como las imágenes en color, que tienen tres componentes). Las funciones tienen un dominio y un rango. Si el dominio y rango son continuos, la señal es continua o analógica; si el dominio es discreto pero el rango no, la señal también será discreta; y si el dominio y el rango son discretos, como en el caso de las imágenes, la señal es digital.

La función de imagen $f(x,y)$ es digitalizada en la memoria del computador, tanto espacialmente como en amplitud. La digitalización de las coordenadas espaciales (x,y) está asociada al concepto de *muestreo*, mientras que la digitalización de la amplitud al de cuantificación de los *niveles de gris*. Una imagen $f(x,y)$ está por tanto almacenada en una matriz de $N \times M$ elementos¹. Entenderemos que el origen de coordenadas de la imagen es la esquina superior izquierda, el eje x el horizontal y el eje y el vertical.

$$f(x,y) = \begin{bmatrix} f(0,0) & f(1,0) & \cdots & f(N-2,0) & f(N-1,0) \\ f(0,1) & f(1,1) & \cdots & f(N-2,1) & f(N-1,1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f(0, M-2) & f(1, M-2) & \cdots & f(N-2, M-2) & f(N-1, M-2) \\ f(0, M-1) & f(1, M-1) & \cdots & f(N-2, M-1) & f(N-1, M-1) \end{bmatrix}$$

El muestreo es la conversión que sufren las dos dimensiones espaciales de la señal analógica, y que genera la noción de píxel. La imagen en la memoria del computador se comporta como una matriz donde el valor de los elementos es su nivel de gris, y la fila y la columna su posición en el espacio. La cuantificación es la conversión que sufre la amplitud de la señal analógica; así se genera el concepto de nivel de gris o

¹ En general N será igual a M y ambas potencias de 2.

intensidad. Para el caso de tener 256 niveles de gris (0-255), el 0 corresponde a un objeto no iluminado o que absorbe todos los rayos luminosos que inciden sobre él (negro), y el nivel 255 a un objeto muy iluminado o que refleja todos los rayos que inciden sobre él (blanco). En la figura 3.1 viene reflejado el proceso de digitalización de un objeto a varias resoluciones. Al efecto de disminuir la resolución y mantener las dimensiones físicas se le conoce como efecto *tablero de ajedrez*.

En el proceso de digitalización es preciso establecer el valor de N y M así como el número de niveles de gris asignados para cada píxel. Es una práctica común en el proceso de digitalización de imágenes que estas cantidades sean números enteros potencias de dos. El número de bits b requeridos para almacenar la imagen digitalizada viene dado por el producto:

$$b = N M g$$

siendo 2^g el número de niveles de gris.

El concepto de resolución surge asociado al número de muestras y niveles de gris necesarios para tener una buena aproximación de la imagen. Es interesante apreciar el efecto de las variaciones de los parámetros citados sobre la calidad de una misma imagen, y cómo es difícil establecer el valor óptimo de cada uno de ellos para los requerimientos de diferentes aplicaciones. Si la imagen posee detalles de poca resolución, las nociones de cuantificación y muestreo estarán íntimamente ligadas.

En la imagen 3.2 aparece el mismo objeto digitalizado con diversos niveles de gris. Es importante notar que aparentemente no existen diferencias entre las primeras imágenes. Esto es debido a que el ojo humano es incapaz de distinguir más de dos docenas de niveles de gris distintos. Conforme disminuye el número de niveles de gris aparece el efecto conocido como *falsos contornos*.

Una vez definidas las resoluciones puede hablarse del *histograma* de una imagen. El histograma contiene el número de píxeles que tienen el mismo nivel de gris (figura 3.3). Puede entenderse como la probabilidad de que un valor de gris determinado aparezca en la imagen. Tiene la gran ventaja que condensa la información de la imagen en muchos menos valores (256 si cada píxel son 8 bits) pero, aunque en los capítulos siguientes se verá que es una de las herramientas más importantes a la hora de procesar y analizar una imagen, el histograma tiene el gran inconveniente de que por sí solo no aporta ninguna información sobre la localización espacial de los píxeles. Así se tendrán infinidad de imágenes distintas que tienen el mismo histograma. A lo largo del libro se verán otras utilidades del histograma como obtener el número de objetos en una imagen, su área y las diversas texturas.

Un último elemento de la imagen son los *planos de bits* (figura 3.4). Si una imagen tiene 256 niveles de gris, cada uno de los píxeles ocupa un *byte* (8 bits):

Pixel x	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
Pixel x+1	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0

Cada plano de bit es una imagen formada por un determinado bit de cada píxel. La importancia de los planos de bits es que reflejan muy bien la influencia del ruido en la imagen. Así puede observarse como para planos de bits altos se pueden distinguir los distintos objetos que forman la imagen, pero para valores bajos la impresión que dan es una distribución aleatoria de puntos.

3.2 Relaciones entre píxeles: vecindad, conectividad.

Un píxel p de coordenadas (x,y) presenta un total de cuatro vecinos en el plano vertical y horizontal, siendo sus coordenadas:

	$x,y-1$	
$x-1,y$	x,y	$x+1,y$
	$x,y+1$	

Este conjunto de píxeles se denomina vecindad de tipo 4 del píxel p , y se representa por $N4(p)$. Además se puede considerar la existencia de otros cuatro vecinos asociados a las diagonales, cuyas coordenadas son:

$x-1,y-1$		$x+1,y-1$
	x,y	
$x-1,y+1$		$x+1,y+1$

los cuales se representan por $ND(p)$. La suma de los anteriores define los ocho vecinos del píxel p , $N8(p)$.

Mediante el concepto de *conectividad* se quiere expresar que dos píxeles pertenecen al mismo objeto, por lo que está relacionado con el de vecindad. Dos píxeles están conectados si son adyacentes (vecinos) y si sus niveles de gris satisfacen algún criterio de especificación (por ejemplo ser iguales).

Existen tres tipos:

- Conectividad-4. Dos píxeles p y q presentan una conectividad-4 si q pertenece al $N4(p)$.
- Conectividad-8. Dos píxeles p y q presentan una conectividad-8 si q pertenece al $N8(p)$.
- Conectividad- m (conectividad mixta). Dos píxeles p y q presentan una conectividad- m si:
 1. q pertenece a $N4(p)$, o
 2. q pertenece a $ND(p)$ y el conjunto $N4(p) \cap N4(q)$ es el conjunto vacío.

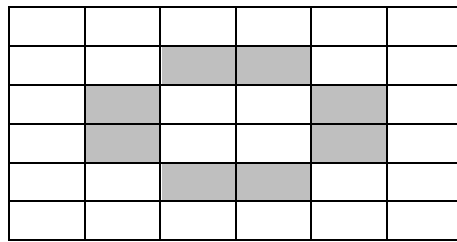
La conectividad mixta es una modificación de la conectividad-8 cuya utilidad es eliminar las conexiones múltiples que en ocasiones aparecen cuando la conectividad-8 es utilizada. Por ejemplo:

1	1	1	1	1	1	1	1	1			
				\ /							
1	-	1	0	1	-	1	0	1	-	1	0
					\						\
0	0	1	0	0	1	0	0	1			
Conectividad-4			Conectividad-8			Conectividad- m					

En este caso, para ir desde el píxel central al inferior derecha se utilizaría conectividad-8. Sin embargo para ir al píxel superior derecha se podría ir de dos formas, directamente o a través del píxel superior. Esta indeterminación se evita con la conectividad mixta. Ahora el píxel superior derecha no está conectado al central ya que sí lo está al superior.

En la figura siguiente se observa que si se utiliza la conectividad de tipo 4, en la figura existen cuatro objetos, cuando de manera intuitiva se observa que es uno solo. Si entonces se toma la de tipo 8 el anterior problema desaparece; pero ahora tanto el fondo

exterior como el interior están conectados. Por ello cuando se quiere determinar si dos pixeles pertenecen al mismo objeto se toma conectividad de tipo 8 y para el fondo conectividad de tipo 4 o viceversa.



3.3 Distancia.

Con la distancia se quiere obtener el mínimo número de pasos elementales que se necesitan para ir de un punto a otro. Dados tres pixeles p, q y z, con coordenadas (x,y), (s,t) y (u,v) respectivamente, se puede definir una función de distancia D si cumple:

- $D(p, q) \geq 0$, ($D(p, q) = 0$, si $p=q$).
- $D(p, q) = D(q, p)$
- $D(p, z) \leq D(p, q) + D(q, z)$

Las funciones de distancia usadas comúnmente son:

La *distancia Euclídea* entre p y q se define:

$$D_E(p, q) = \sqrt{(x - s)^2 + (y - t)^2}$$

con esta definición, las distancias serán:

$$\begin{array}{ccccc} \sqrt{8} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{8} \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ 2 & 1 & 0 & 1 & 2 \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ \sqrt{8} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{8} \end{array}$$

La distancia euclídea es la más exacta pero presenta el inconveniente de su gran número de cálculos. Por otro lado no tiene en cuenta el concepto de vecindad antes descrito. Es por ello que se han definido otras.

Distancia Manhattan. Se toman solamente en cuenta los vecinos de orden 4.

$$D = |x - s| + |y - t|$$

Las distancias serán por tanto:

```

4 3 2 3 4
3 2 1 2 3
2 1 0 1 2
3 2 1 2 3
4 3 2 3 4

```

Con esta definición los vecinos de tipo 4 están a la distancia unidad. Si lo que se quiere es que los vecinos de tipo 8 estén a la misma distancia se toma:

$$D(p,q)=\max(x-s,y-t)$$

Obteniéndose la distancia *tablero de ajedrez*. Y por tanto se tiene:

```

2 2 2 2 2
2 1 1 1 2
2 1 0 1 2
2 1 1 1 2
2 2 2 2 2

```

3.4. Color.

Aunque el color es una de las características más importantes que definen a los objetos, hasta hace poco tiempo se le prestaba poca atención debido al coste computacional y a la memoria necesaria para procesar imágenes en color. Con el desarrollo del hardware esto ha cambiado. Algunas definiciones básicas para comprender los espacios de colores son:

- Brillo: sensación que indica si un área está más o menos iluminada.
- Tono: sensación que indica si un área parece similar al rojo, amarillo, verde o azul o a una proporción de dos de ellos.
- Coloración: sensación por la que un área tiene un mayor o menor tono.
- Luminosidad: brillo de una zona respecto a otra blanca en la imagen.
- Croma. La coloridad de un área respecto al brillo de un blanco de referencia.
- Saturación. La relación ente la coloridad y el brillo.

Los parámetros síquicos de la percepción del color son la luminosidad, el tono, y la saturación. Para obtener una imagen habrá que transformar primero los parámetros cromáticos en eléctricos (análogamente a lo visto para imágenes en blanco y negro) y representar los colores; lo que puede realizarse de distinta manera, dando lugar a los diversos espacios de color. Por tanto un espacio de color es un método por el que se pueda especificar, crear o visualizar cualquier color. Dependiendo del tipo de sensor y aplicación se han desarrollado diversos espacios de colores. Así se tienen espacios de colores que se utilizan para la adquisición, la transmisión de las señales (por ejemplo para la televisión), y la impresión o los espacios que tratan de imitar la percepción humana.

3.4.1. Espacio RGB.

El espacio RGB se basa en la combinación de tres señales de luminancia cromática distinta: el rojo, el verde y el azul (Red, Green, Blue). La manera más sencilla e intuitiva de conseguir un color concreto es determinar la cantidad de color rojo, verde y azul que se necesita combinar (figura 3.5); para ello se realiza la suma aritmética de las componentes:

$$X = R + G + B$$

Gráficamente se representa por un cubo. En la recta que une el origen con el valor máximo se hayan los grises, ya que las tres componentes son iguales. Como se ha comentado con anterioridad, cuando una cámara adquiere una imagen en color, para cada pixel en color se tienen en realidad tres, uno por cada componente; la ganancia máxima de cada uno de ellos corresponde a la longitud de onda de los tres colores básicos antes nombrados (rojo, verde y azul). Aunque el sistema RGB es el más intuitivo de todos y de hecho es en el que se basan las cámaras para adquirir imágenes en color, presenta un serio inconveniente: en sus tres valores mezcla la información del color (tono y saturación) y la intensidad. Así por ejemplo, en la figura 3.6 se tiene una imagen con dos colores dominantes (rojo y verde). Si se quisieran ecualizar los histogramas (técnica que se describirá detalladamente en el capítulo 5) para lograr un mayor contraste en la imagen, se podría tomar cada componente como una imagen en blanco y negro, ecualizarlas y volverlas a juntar. Puede verse que el resultado no es el deseado ya que los colores han cambiado y algunos objetos son ahora azules. Esto ocurre por lo comentado antes de que se tiene la información repartida entre las tres componentes. Por otro lado, si se quisieran separar los objetos por colores, habría que volver a tener en cuenta la información de las tres imágenes. Estos dos inconvenientes se resuelven con el siguiente espacio de color.

3.4.2. El espacio HSI.

El espacio de color HSI (figura 3.7) se basa en el modo de percibir los colores que tenemos los humanos. Dicho sistema caracteriza el color en términos de tono o tinte (Hue), saturación o cromatismo (Saturation) y brillo (Intensity); componentes que se muestran favorables de cara a realizar segmentaciones de la imagen en atención al tono o tinte del color. Así, de dos fuentes de luz con mismo espectro, aquella que tenga mayor intensidad aparecerá más brillante. Existen sin embargo ejemplos de objetos de igual intensidad que no tienen el mismo brillo: una luz verde y otra roja se distinguen por el tono que representa la longitud de onda de ese color. Sin embargo no es igual a la longitud de onda ya que algunos colores que se encuentran en la naturaleza (el púrpura por ejemplo) no se encuentran en la descomposición de la luz blanca. La saturación distingue la blancura de un color (el rojo tiene saturación máxima y el rosa mínima).

Las transformaciones matemáticas que permiten el paso del espacio RGB al HSI son realizadas normalmente mediante hardware específico. Las ecuaciones son:

$$I = \frac{R + G + B}{3}$$

$$H = \arctan\left(\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)}\right)$$

$$S = 1 - \frac{\min(R, G, B)}{I}$$

En la expresión del tono $H = \arctan\left(\frac{x}{y}\right)$, se emplea el signo de x e y para establecer el cuadrante al que pertenece el ángulo resultante. Se puede considerar que el tono indica el ángulo formado entre el eje de referencia (el correspondiente al color rojo) y el punto que ocuparía el color a analizar.

Así como el espacio RGB se representa por un cubo, el HSI lo forman dos pirámides unidas por su base. Dependiendo de la intensidad se tiene un corte a los conos. Dentro de la circunferencia obtenida el color viene definido por su orientación (el rojo es 0°, el verde 120° y el azul 240°) y la saturación indica la lejanía del centro del círculo.

Variaciones de este espacio de color son:

- HSL (Hue Saturation Lightness)
- HSV (Hue Saturation Value)
- HCI (Hue Chroma / Colourfulness Lightness)
- HVC (Hue Value Chroma)
- TSD (Hue Saturation Darkness)

Otros autores prefieren utilizar las siguientes transformaciones matemáticas del espacio de color RGB a HSI.

$$H = \arccos\left(\frac{0.5((R - G) + (R - B))}{\sqrt{(R - G)^2 + (R - B)(G - B)}}\right)$$

Si $(B/I) > (G/I)$ entonces $H = 360 - H$.

o también:

$$H = \frac{\alpha - \arctan\left(\frac{\sqrt{3}(R - I)}{G - B}\right)}{2\pi}$$

con:

$$\alpha = \frac{\pi}{2} \text{ si } G > B$$

$$\alpha = \frac{3\pi}{2} \text{ si } G < B$$

En el caso de tener un color con igual componente verde y azul ($G=B$), el tono será igual a la unidad ($H=1$).

o también:

$$S = \sqrt{R^2 + G^2 + B^2 - 2RG - RB - BG}$$

La conversión inversa, es decir, aquella que permite recuperar de nuevo la información de la imagen en el espacio RGB, viene dada por las siguientes relaciones en función de los valores tomados por el tono:

Si $0^\circ < H \leq 120^\circ$

$$B = \frac{1}{3}(1 - S)$$

$$R = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$G = 1 - (B + R)$$

$120^\circ < H \leq 240^\circ$

$$H = H - 120^\circ$$

$$R = \frac{1}{3}(1 - S)$$

$$G = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$B = 1 - (R + G)$$

$240^\circ < H \leq 360^\circ$

$$H = H - 240^\circ$$

$$G = \frac{1}{3}(1 - S)$$

$$B = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$R = 1 - (G + B)$$

Los inconvenientes que presenta el uso del espacio HSI son:

- La no linealidad de las transformaciones realizadas, lo que implica un alto coste computacional para obtenerlas.
- La singularidad existente en el entorno del eje definido por $R=G=B$ (saturación=0), que provoca que los valores obtenidos para el tono sean inestables.

Las ventajas se aprecian en la figura 3.8. La componente H describe el tono de los objetos, por lo que si ahora obtenemos su histograma, se aprecian dos grupos de píxeles en torno a los valores rojos y verdes. Por otro lado, si queremos modificar la luminosidad de la imagen, se tomaría solamente la componente I, se modificaría y se volvería a componer la imagen, que como se observa no cambia de colores.

3.4.3. Los espacios XYZ, Luv, Lab.

Con las representaciones anteriores del espacio de color aparecen coeficientes negativos para colores de determinada longitud de onda, lo cual supone una fuente de error importante en los cálculos colorimétricos (comparación de un color monocromático con el blanco de referencia). La CIE (Commission Internationale de l'Eclairage) estableció un nuevo espacio de color, el XYZ, con la finalidad de evitar esos coeficientes negativos. En la componente Y estaría la luminosidad y en XZ la coloración; el eje Y es perpendicular al plano definido por XZ. Este espacio es independiente del dispositivo que se esté usando. Se suele trabajar con valores normalizados entre cero y uno. Así:

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z}$$

y para una determinada luminosidad Y:

$$X = \frac{xY}{y} \quad Z = \frac{zY}{y}$$

Para pasar de RGB a XYZ existen varias normativas distintas, la más reciente de todas es la de la ITU (International Telecommunications Union).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.431 & 0.342 & 0.178 \\ 0.222 & 0.707 & 0.071 \\ 0.020 & 0.130 & 0.939 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.063 & -1.393 & -0.476 \\ -0.969 & 1.876 & 0.042 \\ 0.068 & -0.229 & 1.069 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

El espacio de colores XYZ obtenido constituye el denominado “Diagrama Internacional xy, CIE” o carta cromática xy. Una vez obtenidas las coordenadas XYZ se pueden construir diferentes espacios CIEs. Entre ellos destacan el CIE L*u*v* y el CIE L*a*b*. En ambos la información se descompone en tono, cromatismo e intensidad de color.

El espacio CIE L*u*v* se basa en las siguientes relaciones donde, de nuevo, L es la información de la luminosidad y uv la del color:

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \quad \text{si } \frac{Y}{Y_n} > 0.008856$$

$$L^* = 903.3 \frac{Y}{Y_n} \quad \text{si } \frac{Y}{Y_n} < 0.008856$$

$$u^* = 13L^*(u' - u'_n)$$

$$v^* = 13L^*(v' - v'_n)$$

siendo:

$$u' = \frac{4X}{(X + 15Y + 3Z)}$$

$$v' = \frac{9Y}{(X + 15Y + 3Z)}$$

y u'_n, v'_n tienen las mismas ecuaciones y son los valores para la referencia del blanco.

En el espacio CIE L*a*b* (figura 3.9), L es la luminosidad, a la referencia respecto a la relación rojo/verde y b respecto a la relación amarillo/azul. Las ecuaciones que lo definen dependen de la relación $\frac{Y}{Y_n}$ y son:

Si $\frac{Y}{Y_n} > 0.008856$

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16$$

$$a^* = 500 \left(\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right)$$

$$b^* = 200 \left(\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right)$$

Si $\frac{Y}{Y_n} \leq 0.008856$

$$L^* = 903.3 \frac{Y}{Y_n}$$

$$a^* = 7.87 \left(\frac{X}{X_n} - \frac{Y}{Y_n} \right)$$

$$b^* = 7.87 \left(\frac{Y}{Y_n} - \frac{Z}{Z_n} \right)$$

El cromatismo en este espacio queda definido por:

$$croma = \sqrt{(a^*)^2 + (b^*)^2}$$

Mientras que el tono o tinte:

$$\text{tono} = \arctan \frac{b^*}{a^*}$$

En la figura 3.10 puede verse la descomposición de la imagen en las tres componentes. La componente de la luminosidad L difiere levemente de la obtenida en el espacio HSI. Como la componente a indica la relación rojo/verde, puede verse en su histograma que la separación entre los grupos de píxeles es mayor que en la componente H, por lo que se podrían separar más fácilmente.

3.4.4. El espacio YIQ, YUV, YCbCr, YCC.

Estos espacios se basan en la obtención de la luminancia (luminosidad) y la crominancia (color). Están muy relacionados con los sistemas de color para la televisión: YIQ para el NTSC, YUV para el PAL e YCbCr para la televisión digital. Son los sistemas utilizados en algunas cámaras de color.

A partir de la elección de los colores primarios y del nivel de blanco de referencia se puede obtener la expresión fundamental de la luminancia monocroma como:

$$Y = 0.30 R + 0.59 G + 0.11 B$$

Siendo RGB las señales de salida de la cámara de vídeo tricolor.

En cuanto a la crominancia, se han experimentado diversas formas de señal; inicialmente se trabajó con las señales R y B de forma aislada, pero al no satisfacer el criterio de que la información debe desaparecer cuando el color es acromático en la escena (color gris), y cuando se transmite en monocromo (televisión en blanco y negro), se buscaron otras formas de señal que son las llamadas señales de diferencia de color.

Las señales simples de crominancia se obtienen al restar a las señales primarias la señal de luminancia Y. Para transmitir la información de luminancia y crominancia se tendrían en total cuatro señales; bastan tres de ellas, al igual que en el caso RGB. La señal que se elimina es G-Y ya que es una señal de baja relación debido a sus menores coeficientes respecto a las señales R-Y y B-Y.

Las ecuaciones son:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

o también:

$$\begin{aligned} U &= 0.493 (B - Y) \\ V &= 0.877 (R - Y) \end{aligned}$$

Por lo que U y V pueden tomar valores positivos o negativos.

La reducción en amplitud mostrada en estas dos últimas expresiones se debe a que, si se modula una portadora directamente con las señales $U=(B - Y)$ y $V=(R - Y)$, se produce una sobremodulación que provoca una distorsión en los colores. La relación anterior permite que la relación señal ruido sea la mejor posible. Así se reduce además el ancho de banda necesario para la transmisión. La componente Y tiene un ancho de banda de 5 MHz y U y V de 2.5 MHz.

La transformación inversa viene dada por:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 & 1.140 \\ 1.000 & -0.396 & -0.581 \\ 1.000 & 2.029 & 0.000 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

El sistema NTSC usa las componentes YIQ (figura 3.11) que, de forma similar a la anterior, representan luminosidad y color. La señal Y tiene un ancho de banda de 4.2MHz, Y 0.5MHz y Q 1.5MHz. El sistema de referencia forma un ángulo de 33° respecto al YUV. Las ecuaciones son:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.569 & -0.274 & -0.322 \\ 0.211 & -0.523 & -0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

La transformación inversa:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & -1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

Si se aplican las fórmulas clásicas de transformación de ejes de coordenadas resultan las siguientes relaciones:

$$I = -0.27(B - Y) + 0.74(R - Y)$$

$$Q = 0.41(B - Y) + 0.48(R - Y)$$

Y por lo tanto:

$$I = 0.6 R - 0.28 G - 0.32 B$$

$$Q = 0.21 R - 0.52 G + 0.31 B$$

Recíprocamente se tiene:

$$R - Y = 0.96 I + 0.62 Q$$

$$B - Y = -1.10 I + 1.70 Q$$

3.4.5. El espacio CMY(K)

Está formado por Cían, Magenta, Amarillo y Negro. Se usa para impresión y fotografías. Al igual que el espacio RGB, el CMY(K) depende del dispositivo (la impresora). Mientras que en RGB se obtiene un color con la suma de sus tres componentes, en CMY(K) se logra con la resta de las componentes al color blanco. Para pasar de RGB a este sistema:

$$C = 1 - R$$

$$M = 1 - G$$

$$Y = 1 - B$$

y al contrario:

$$R = 1 - C$$

$$G = 1 - M$$

$$B = 1 - Y$$

Si se usa la componente negra:

CMY a CMYK

$$B = \text{minimo}(C, M, Y)$$

$$C = \frac{C - B}{1 - B}$$

$$M = \frac{M - B}{1 - B}$$

$$Y = \frac{Y - B}{1 - B}$$

y la inversa

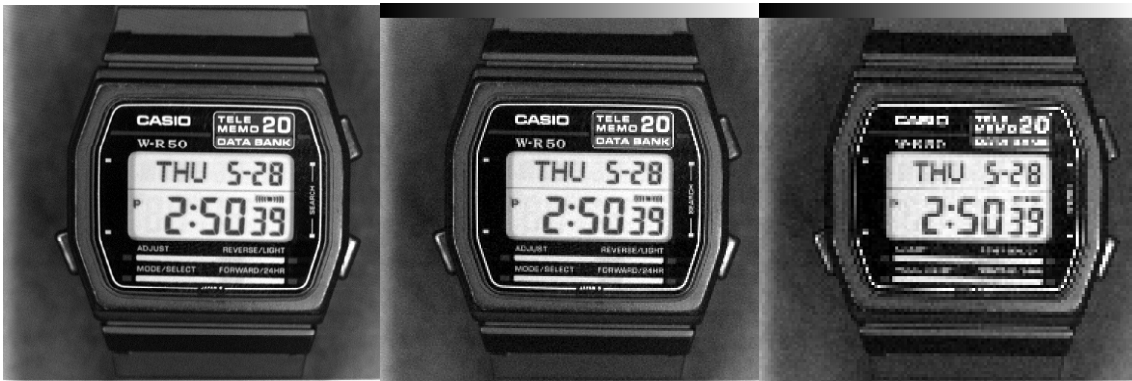
$$C = \text{mínimo}(1, C(1 - B) + B)$$

$$M = \text{mínimo}(1, M(1 - B) + B)$$

$$Y = \text{mínimo}(1, Y(1 - B) + B)$$

3.5 Bibliografía.

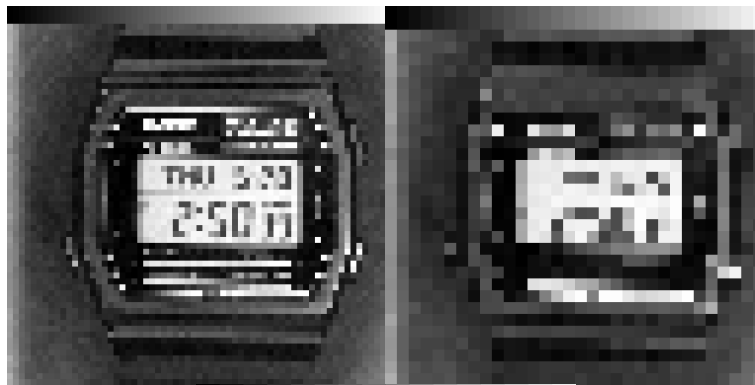
- González, R. C.; Woods, R E.. *Digital image processing*, Addison-Wesley, 1993
- Pratt, W. K. *Digital image processing*,. 2ª edición. John Wiley & Sons, cop. 1991.
- Jain, A. K. *Fundamentals of digital image processing*, Prentice-Hall International 1989
- Ohta, Y., Kanade, T., Sakai, T. Color information for region segmentation. *Computer Graphics and Image Processing* 13, 222-241,1980.
- Russ, J. C. *The image processing handbook*,. 2ª edición. CRC Press, 1995



512x512

256x256

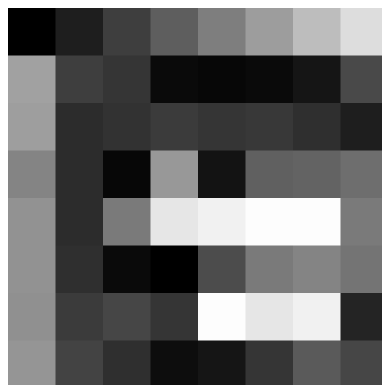
128x128



64x64

32x32

16x16



8x8

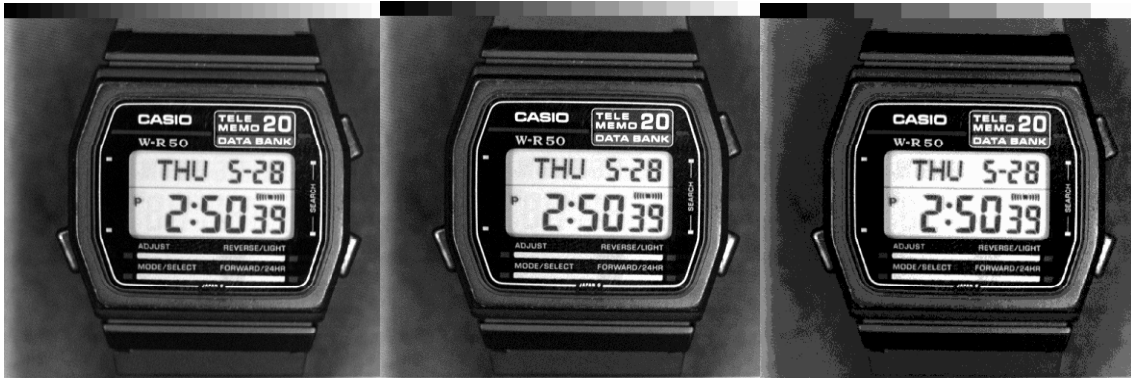
Figura 3.1 Diversos muestreos espaciales



256

128

64



32

16

8



4

2

Figura 3.2 Diversos niveles de gris.

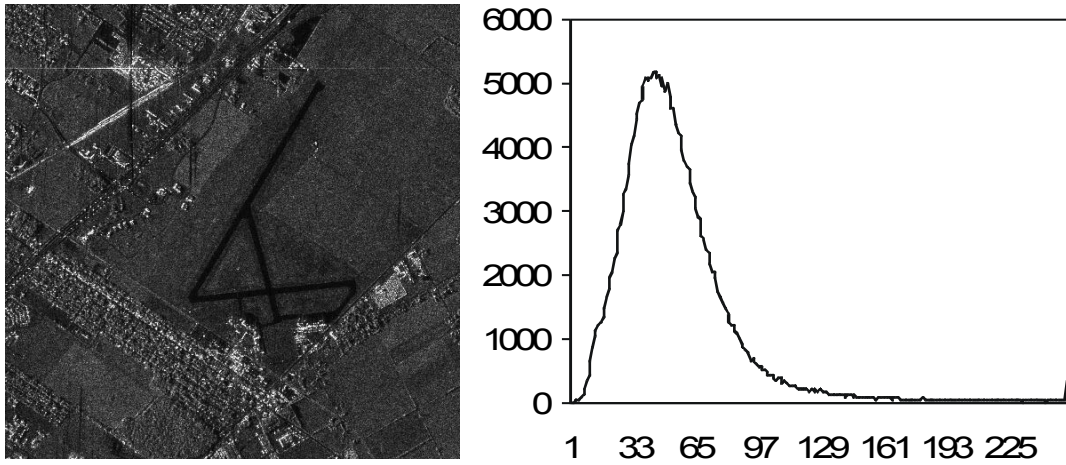


Figura 3.3. Imagen aérea de un aeropuerto y su histograma.

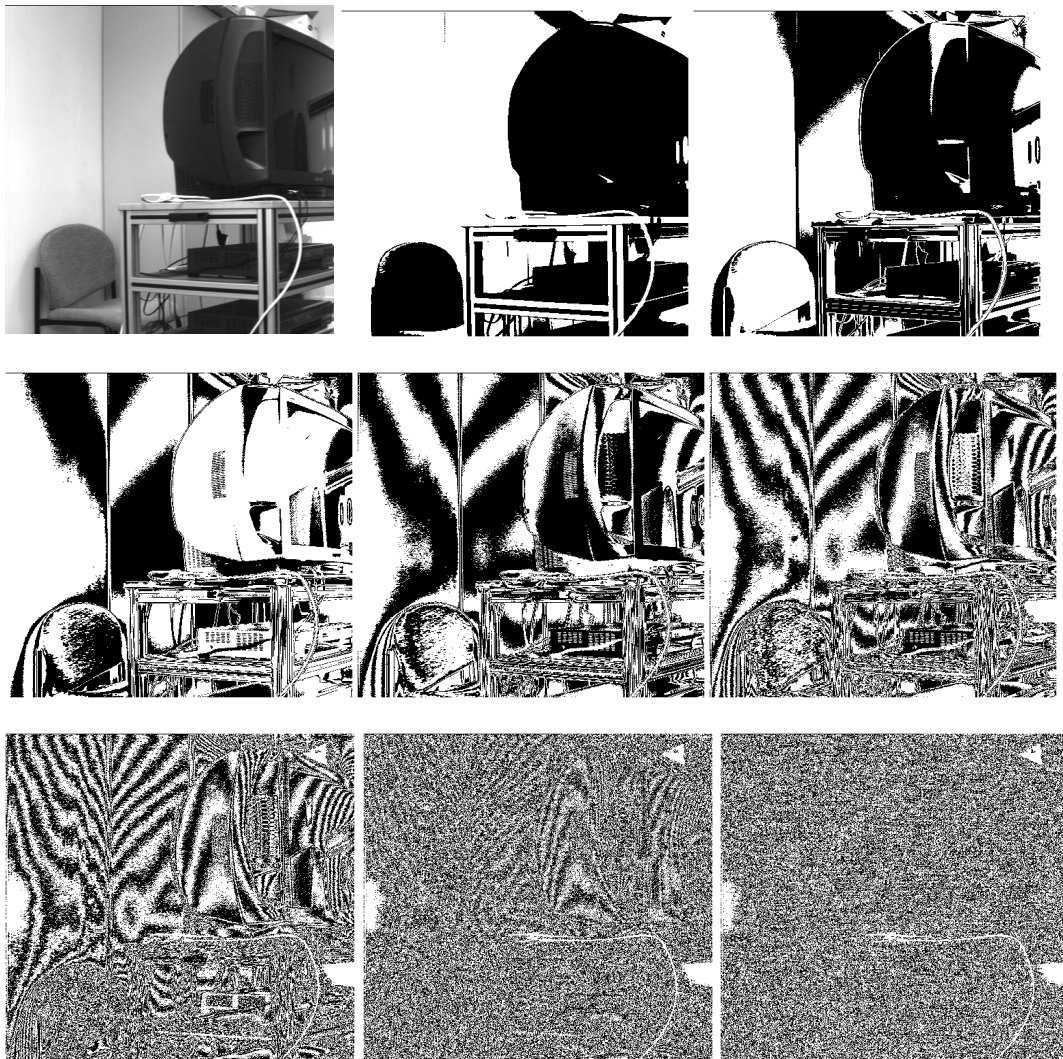


Figura 3.4. planos de bits

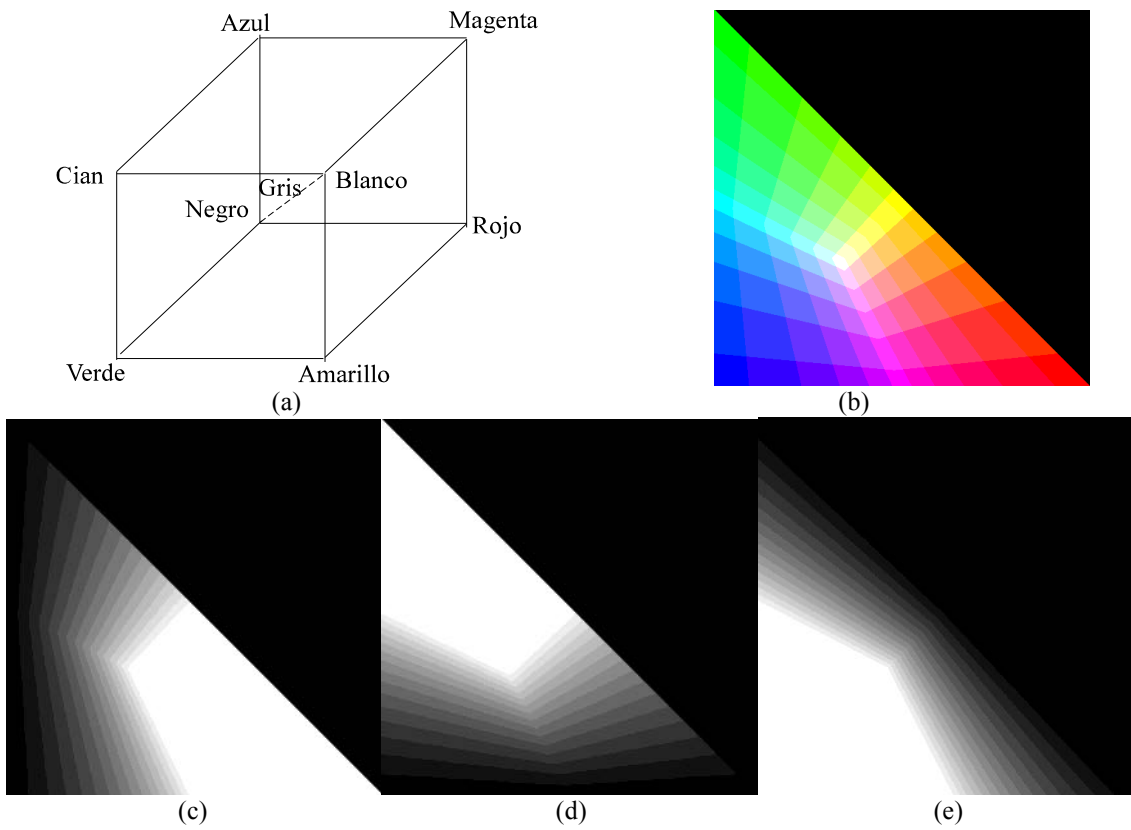


Figura 3.5 Espacio de colores RGB. (a) Espacio RGB (b) Imagen (c) Componente roja (d) Componente verde (e) Componente azul

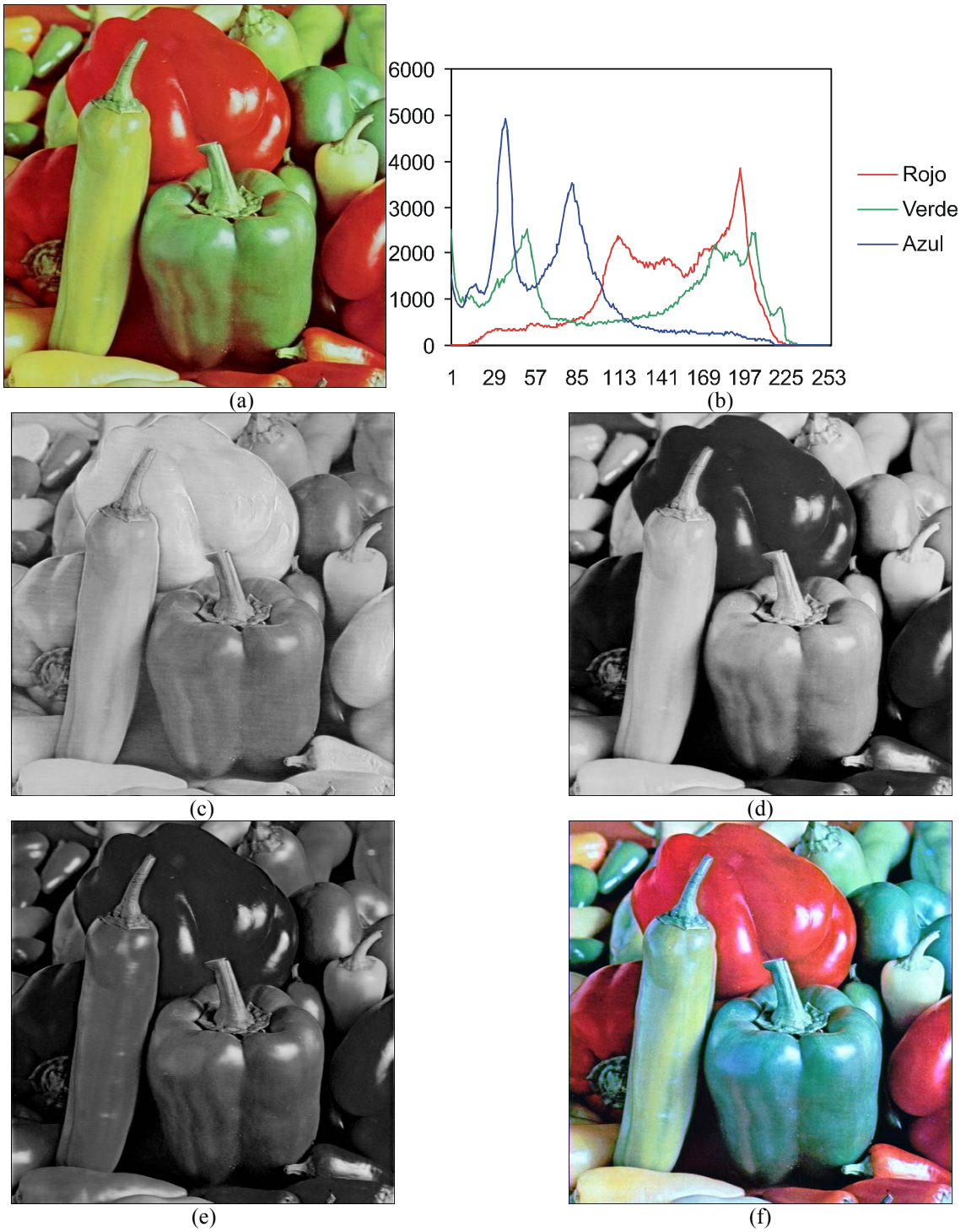


Figura 3.6 Análisis de una imagen RGB. (a) Imagen original (b) histogramas de cada componente (c) componente roja (d) componente verde (e) componente azul (f) resultado de equalizar las tres componentes.

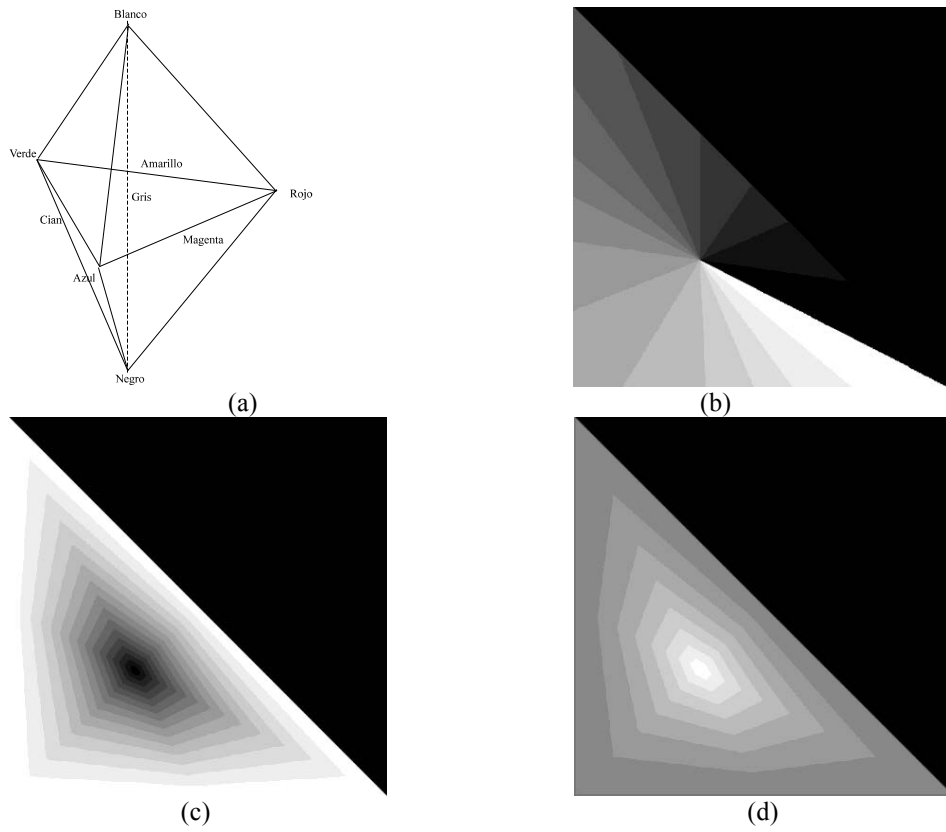


figura 3.7 Espacio de colores HSI (a) Espacio de colores HSI (b) Componente H (c) Componente S (d) Componente I

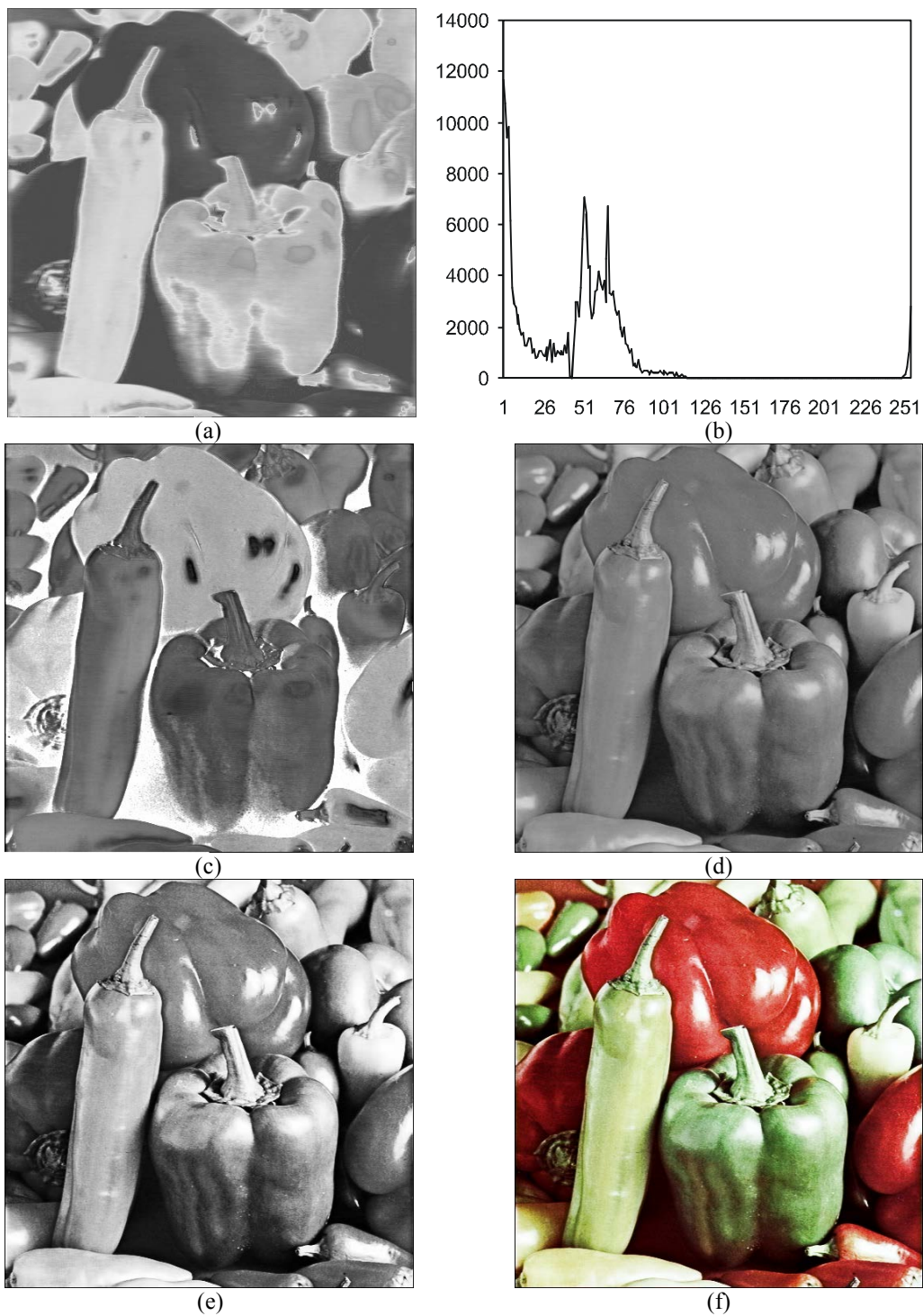


Figura 3.8 Análisis de una imagen HSI. (a) componente H (b) histogramas de la componente H (c) componente S (d) componente I (e) ecualización de la componente I (f) nueva imagen en color.

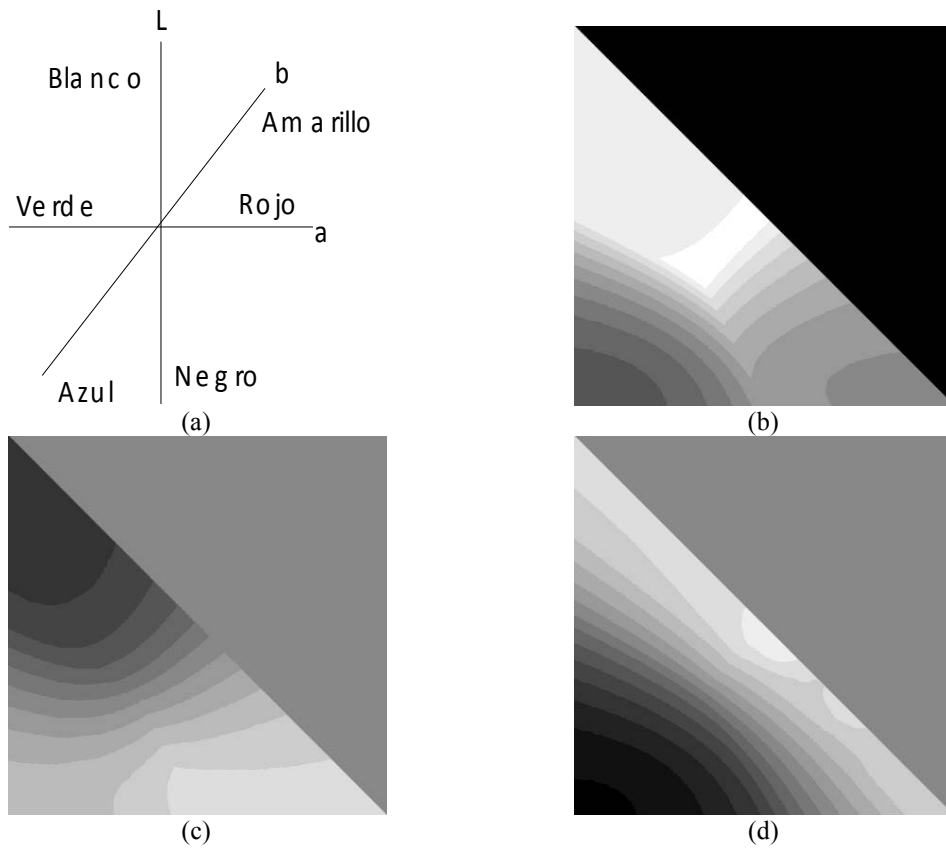


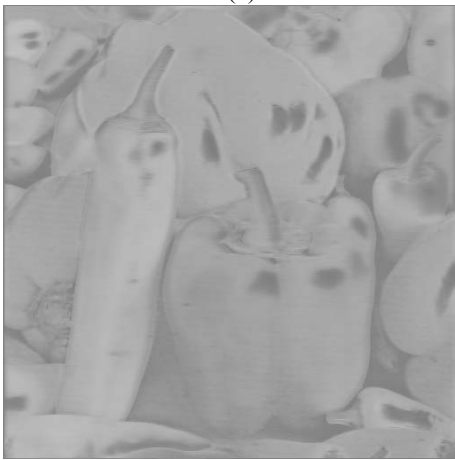
Figura 3.9 Espacio de colores $L^*a^*b^*$. (a) Espacio de colores (b) componente L^* (c) componente a^* (d) componente b^*



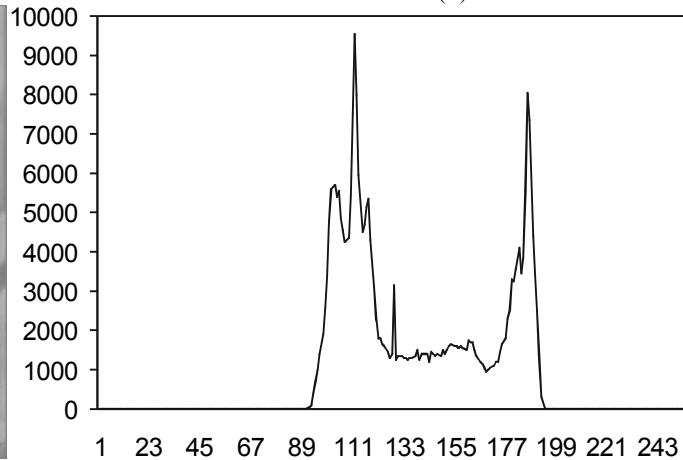
(a)



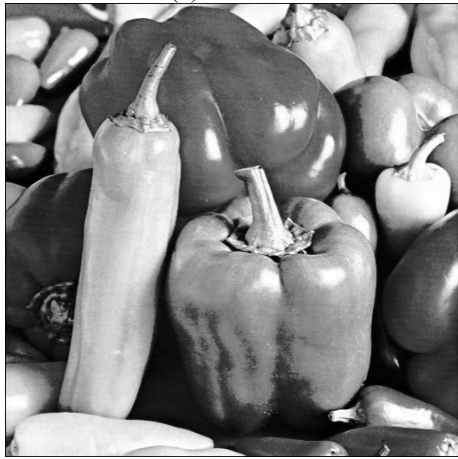
(b)



(c)



(d)



(e)



(f)

Figura 3.10 Análisis de una imagen $L^*a^*b^*$. (a) componente L^* (b) componente a^* (c) componente b^* (d) histogramas de la componente a^* (e) ecualización de la componente L^* (f) nueva imagen color.

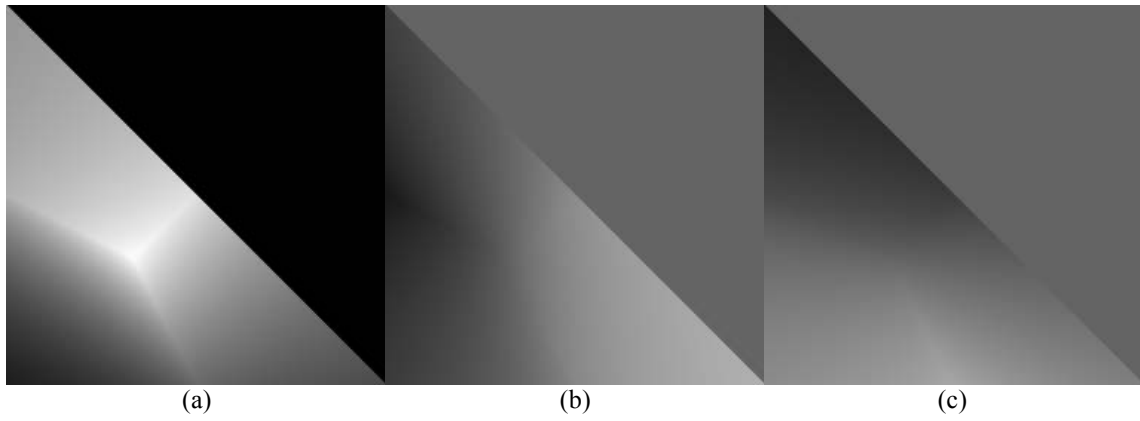


figura 3.11 Espacio de colores YIQ. (a) componente Y (b) componente I (c) componente Q

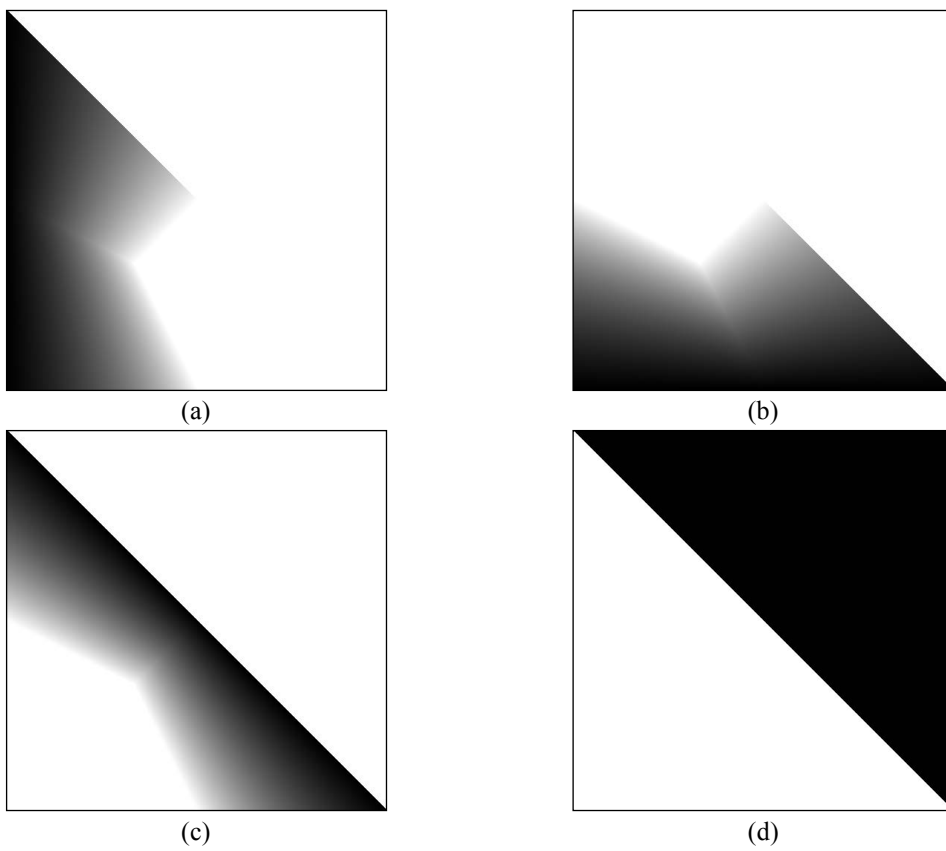


figura 3.12 Espacio de colores CMYK. (a) componente C (b) componente M (c) componente Y (d) componente K