Procesamiento y Análisis de Imágenes

Profesor Gamaliel Moreno

Universidad Autónoma de Zacatecas
Grupo de investigación PDS-UAZ
2025

Contenido

1 Introducción

1.1 Ciencia de imágenes

La ciencia de imágenes puede dividirse en:

- Adquisición. Entrada un magnitud física y salida una imagen.
- Procesamiento. Entrada y salida una imagen.
- Análisis. Entrada una imagen, salida estructuras de datos.
- Comprensión. Entrada una imagen y salida estructuras semánticas
- Visión computacional. Entrada una imagen y salida tomada de decisiones.
- Visión industrial. Entrada una imagen y salida señales de control.

Otra manera de clasificar la ciencia de imágenes es por sus tareas:

- Mejoramiento de la imagen
  - Eliminación o reducción de ruido,
  - Mejorar de contraste, nitidez, enfoque,
  - Datos faltantes (inpainting)
  - Incrementar resolución.

- Extracción de características
  - Segmentación,
  - detección,
  - reconocimiento,
  - registro.

o

1.2 Definición de imagen

La definición conceptual de imagen es la imitación o reproducción de una escena de la realidad mediante un sistema perceptivo. Una imagen es digital cuando se discretizan, cuantizan sus componentes espaciales y de intensidad y esta ultima se codifica. Intuitivamente, una imagen la definimos como una rejilla rectangular (2D) de pixeles con tonos, en la figura 1 se muestra un ejemplo.

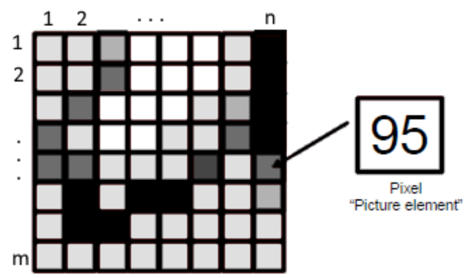

Figure 1: Ilustración de una imagen bidimensional en escala grises.

Para poder analizar las imágenes es necesario formalizar su definición. El caso más sencillo es una imagen bidimensional como una función, su formalización matemáticamente se escribe en la definición 1.

Definición 1. *Una imagen I puede definirse como un mapeo de un pixel x perteneciente a una rejilla $R = \{x \in (i \cdot h \times j \cdot h)|i, j \in N\}$ a un espacio de estado E. Así $I : \mathbb{R}^2 \to \mathbb{R}^c$ es una función acotada inferior y superiormente, donde c representa las características o canales. Si $h = 1$ entonces $i = \{1, \ldots, m\}$, $j = \{1, \ldots, n\}$ y el tamaño de la imagen será será $m \times n$.*

En ciertas aplicaciones las imágenes no son bidimensionales sino tridimensionales, como el caso de las imágenes de resonancia magnética, incluso puede definirse de manera abstracta en altas dimensiones, véase la definición 2.

Definición 2. *Una imagen es un mapeo de un pixel $\mathbb{X}$ de la rejilla regular $R = \{\mathbb{X} \in (\mathbb{X}_1 \times \mathbb{X}_2 \cdots \times \mathbb{X}_d)|\mathbb{X}_1, \mathbb{X}_2, \ldots \mathbb{X}_d \in \mathbb{N}\}$ a un espacio de estado E. Así $I : \mathbb{R}^d \to \mathbb{R}^c$. Si $d = 2$ se habla de pixeles, si $d = 3$ de voxels y $d > 3$ de hipervoxels. El tamaño de la imagen es el producto de la cardinalidad de los conjuntos $\prod_{j=1}^{d} \#\mathbb{X}_j$.*

A partir de la definición 2 tenemos la imagen típica bidimensional ($d = 2$), escala grises ($c = 1$), de 8 bits $Rango(I(x)) \in \{(0, 255)\}$. Algunas variaciones de lo anterior tiene un nombre específico, por ejemplo; imagen binaria con $d = 2$, $c = 1$ y $Rango(I(x)) \in \{0, 1\}$, imagen de color de 16 bits con $d = 2$, $c = 3$ y $Rango(I(x)) \in (0, 65535)$, imagen RGB-D con $d = 2$, $c = 4$, con diversos bits.

Las imágenes pueden tener una definición formal como matriz, como conjunto o grafo. Aquí mostramos la dos primeras en la definición 3 y 4 respectivamente.

Definición 3. *Una imagen bidimensional es una matriz que asigna un número real o natural a cada par de índices (j, i) donde j representa las filas e i las columnas*

$$I = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1n} \\ I_{21} & I_{22} & \cdots & I_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ I_{m1} & I_{m2} & \cdots & I_{mn} \end{bmatrix}.$$

Si $I : \mathbb{R}^d \to \mathbb{R}^c$ se tiene una imagen en espacio discreto, y si $I : \mathbb{R}^d \to \mathbb{N}^c$ se dice que I es una imagen digital.

Definición 4. *Una imagen esta definida* $I = \{\boldsymbol{p} | \boldsymbol{p} = \langle \boldsymbol{x}, \boldsymbol{c} \rangle, x \in \mathbb{X}, c = I(\boldsymbol{x})\}$ *donde* p *son los pixeles que están formados por una tupla* $\langle \boldsymbol{x}, \boldsymbol{c} \rangle$ *donde* $\boldsymbol{x}$ *es un vector de posición y* $\mathbf{c}$ *un vector de características. Dos operadores importantes son*

$$pos(\boldsymbol{p}) = pos(\langle \boldsymbol{x}, \boldsymbol{c} \rangle) = \boldsymbol{x},$$

$$val(\boldsymbol{p}) = pos(\langle \boldsymbol{x}, \boldsymbol{c} \rangle) = \boldsymbol{c}.$$

1.3 Mejoramiento

Definición 5. *El mejoramiento consiste en dada una imagen deteriorada* $f \in \mathsf{I}$ *encontrar un procesamiento* $T : \mathsf{I} \to \mathsf{I}$ *tal que* $g := T(f)$ *la imagen sea restaurada. Una definición más general de procesamiento es: Un procesamiento* $\mathscr{O} : \mathsf{I} \to \mathsf{I}$ *cambia una imagen* $f \in \mathsf{I}$ *en otra imagen* $g \in \mathsf{I}$*.*

El procesamiento puede considerar solo la información del pixel a tratar, considerar todos los elementos de la imagen, tomar información de un vecindario local, un vecindario no local o información externa a la imagen. Inicialmente trataremos con información del pixel, de toda la imagen y un vecindario local.

1.4 Vecindario local

Definición 6. *Por simplicidad tomaremos el caso bidimensional. Un vecindario local de un pixel* x_i *esta dado por* $N_h(x_i) = \{x : a \leq L_p(x_i, x) \leq b; a, b \in \mathbb{R}^+\}$ *donde* L_p *es la distancia de Minskowsky entre el pixel* x_i *y* x *véase la figura 2.*

Veamos la definición de dos vecindarios: el vecindario 4 dado por $N_4(x_i) = \{x : L_1(x_i, x) = 1\}$ y el vecindario 8 definido como $N_8(x_i) = \{x : 1 \leq L_2(x_i, x) \leq \sqrt{2}\}$

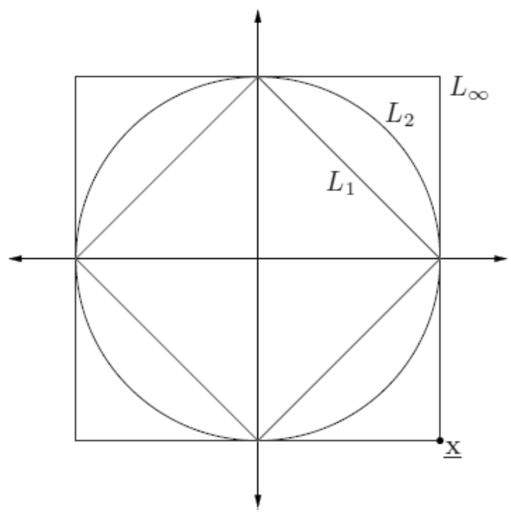

Figure 2: Distancia de Minkowsky en un plano $\mathbb{R}^2$ del origen a 1.

Para simplificar, la notación cuando utilicemos el vecindario en operaciones este incorporará el pixel en el que se esta operando, es decir, el vecindario será $N_h(x_i) \cup x_i$.

1.5 Operaciones

Las operaciones se pueden categorizar como: operaciones sobre píxeles, sobre vecindarios y geométricas. Si las imágenes son definidas como funciones las operaciones binarias de las imágenes f y g de suma $f(x) + g(x)$, resta $f(x) - g(x$, multiplicación $f(x) \times g(x)$ y división $f(x)/g(x)$ se hacen pixel por pixel. Se debe tener cuidado si se define la imagen como matriz ya que la multiplicación y división no se definen de las misma manera.

Si la imagen es definida como conjunto se pueden realizar operaciones de conjunto como unión $(f \cup g)$, intersección $(f \cap g)$, complemento (f^c), diferencia (f/g) entre otras. La definición de la imagen como conjunto es útil cuando se utilizan operaciones morfológicas y de segmentación.

Generalizando las operaciones tenemos:

- La operación en una imagen F sobre los valores de los píxeles se define como

$$G = \{g : g = \langle \boldsymbol{x}, \boldsymbol{d} \rangle, \quad \boldsymbol{d} = T(\boldsymbol{c}), \quad f = \langle \boldsymbol{x}, \boldsymbol{c} \rangle \in \mathsf{F}\}$$

  lo que se interpreta como una modificación del valor de cada píxel.

- La operación sobre una vecindad

$$G = \{g : g = \langle \boldsymbol{x}, \boldsymbol{d} \rangle, \quad \boldsymbol{d} = T(N(\boldsymbol{x})), \quad f = \langle \boldsymbol{x}, \boldsymbol{c} \rangle \in \mathsf{F}\}$$

- La operación geométrica

$$G = \{g : g = \langle \boldsymbol{y}, \boldsymbol{c} \rangle, \quad \boldsymbol{y} = T(\boldsymbol{x}), \quad f = \langle \boldsymbol{x}, \boldsymbol{c} \rangle \in \mathsf{F}\}.$$

  La imagen cambia su codominio sin modificar su rango.

En las operaciones geométricas hay tres importantes: similares, afines y proyectivas, las cuales son lineales. Estas suelen hacer rotación, escalamiento isotrópico y traslación.

2 Contraste e histograma

Este procesamiento se realiza en el domino espacial, se le llama procesamiento en el espacio para diferenciarlo con el procesamiento en otros dominios como el frecuencial. Iniciaremos con procesamiento de mejora.

2.1 Histograma

La mejorar de contraste y nitidez consiste en transformación de pixeles. Transforma una imagen $f(x, y)$ en otra $g(x, y) = T(f(x, y))$ por medio de la función T. Antes de mostrar algunas funciones de contraste es necesario definir el histrograma de imagen, véase la definición 7 y la figura 3.

Definición 7. *Histograma. Sea $R = [0, L-1]$ el rango de la imagen I, el cual será K-particionado exhaustivamente, esto es, $R = [r_0, r_1) \cup [r_1, r_2) \cup \cdots [r_k, r_{k+1}) \cup \cdots [r_{K-1}, r_K]$. Entonces, el histograma $h(k) = n_k$ es el número de pixeles (n_k) asociado al k-ésimo subintervalo. Denotado como conjunto tenemos $h(k) = \#\{x | x \in [r_k, r_{k+1})\}$.*

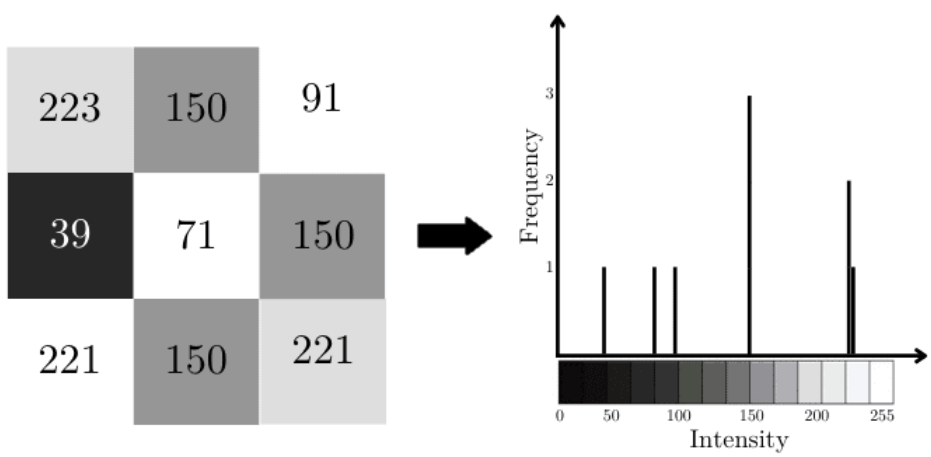

Figure 3: Histograma. A la izquierda una imagen escala grises de 8 bits de tamaño 3×3, a la derecha su histograma.

Si hacemos lo mostrado en la ecuación 1 el podemos pasar del histograma a una función de masa de probabilidad, dado que se cumplen con los tres axiomas de la medida de probabilidad.

$$p(k) = \frac{h(k)}{N}, \text{ con } N = \sum_{k=0}^{K-1} h(k) \tag{1}$$

Tabla 1: Transformaciones básicas

Transformación	Modelo
Normalización	$s = \frac{r}{L-1}$
Logarítmica	$s = c * log(1 + r)$ [a]
Exponencial	$s = e^{r/c} - 1$
Corrección gamma	$s = cr^{\gamma}$

[a]donde c es una constante de normalización que acota los tonos

2.2 Mejora de contraste

Una vez definido el histograma podemos decir que una imagen I de $L = 2^{n-bits}$ tiene mal contraste cuando $min(I) > 0 \vee máx(I) < L$, es decir los tonos no cubren todo el rango de tonos posibles, es común nombrarle histograma estrecho.

Supongamos una imagen en escala grises de 8 bits con un histograma estrecho como el mostrado en la figura 4a. La función de transformación es la expresión de una recta $y(x) = mx + b$ cuyos parámetros están en términos de mínimo y máximo de valor de la imagen de entrada, si definimos a r un pixel de entrada y s el pixel transformado o de salida tenemos

$$s = (L - 1)\frac{r - min(I)}{máx(I) - min(I)} \tag{2}$$

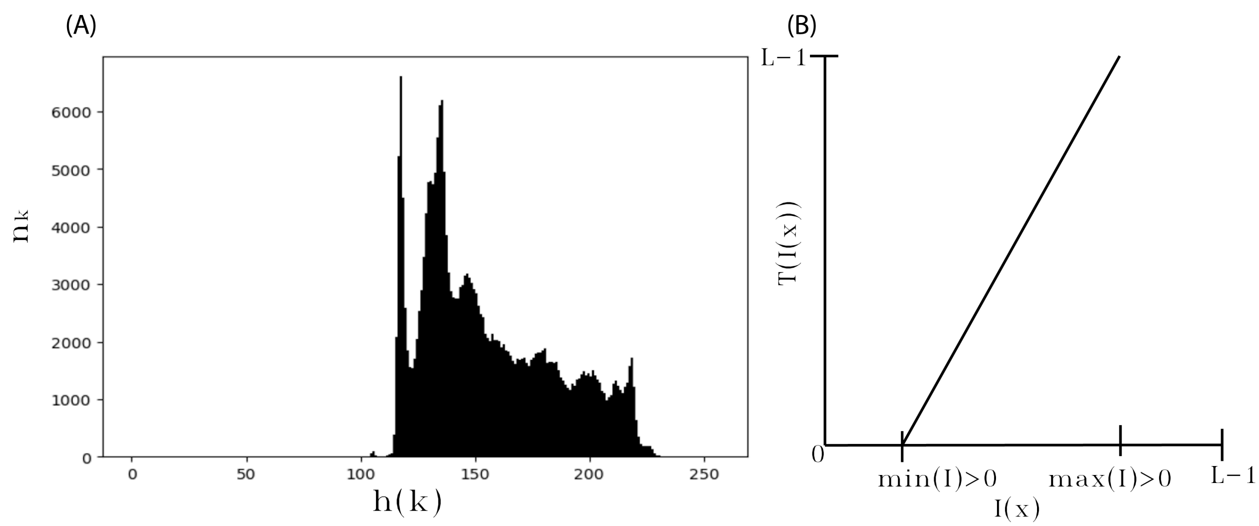

Figure 4: En (A) ejemplo de un histograma estrecho de una imagen escala grises de 8-bits y en (B) una transformación lineal que mejora el contraste.

Existen una gran diversidad de funciones de transformación para mejorar el contraste de una imagen, en la Figura 5 se muestran las gráficas de algunas de ellas y en la Tabla 1 se muestran las expresiones.

2.3 Ecualización de histograma

En la mejora de contraste, desde un enfoque probabilístico, se busca que el histograma tenga una distribución uniforme. Esto se puede lograr utilizando la transformada integral de probabilidad.

Definición 8. *La transformada integral de probabilidad establece que si X es una variable aleatoria continua con distribución acumulada F_X, entonces la variable aleatoria $Y = F_X(X)$ tiene una distribución uniforme en $U \sim Unif(0, 1)$.*

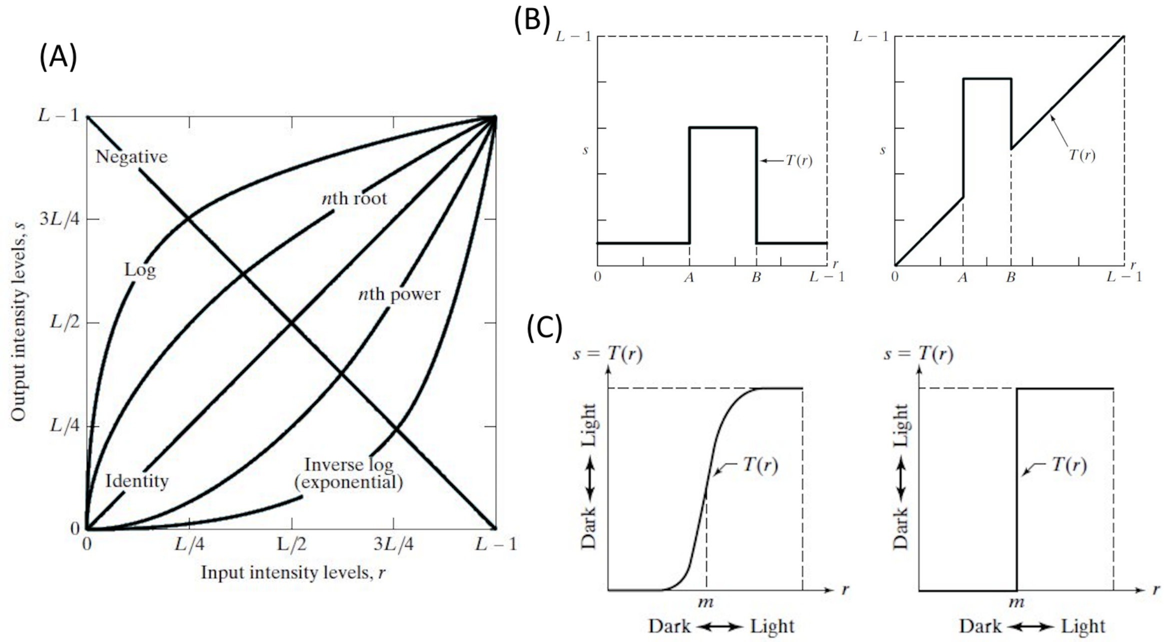

Figure 5: En (A) funciones típicas de transformación para el contraste, en (B) función de transformación por intervalos y en (C) función sigmoide.

Prueba. Sea Y una variable aleatoria definida por $Y = F_X(X)$ donde X es otra variable aleatoria. Entonces

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(F_X(X) \leq y) \\ &= P(F_X(X) \leq F_X^{-1}(y)) \\ &= F_X(F_X^{-1}(y)) \\ &= y. \end{aligned}$$

$F_Y(y) = y$ describe la función de distribución acumulada de la variable aleatoria $Y = F_X(X)$, uniformemente distribuida en $[0, 1]$. Enmarquemos esta definición en el contexto de una imagen. Sea I una imagen escala grises con pixeles x con una distribución de masa probabilidad $p(x)$ y un total de escalas L. La probabilidad de ocurrencia de un $k - nivel$ de gris esta dada por la ecuación 3 y su acumulada por la ecuación 4.

$$p_x(k) = p(x = k) = \frac{h(k)}{N}. \tag{3}$$

$$F_x(k) = \sum_{j=0}^{k} p(x = j). \tag{4}$$

Ahora, deseamos una función de transformación que $Y = T(X) \sim U$, esta función es la definida en 8, sin embargo en el contexto de imagen digital la transformación debe estar en un intervalo de $0 \leq k < L$ y en entero, considerando esto la función queda como se muestra en la ecuación 5.

$$Y(k) = \lfloor F_x(k)(L - 1) \rfloor. \tag{5}$$

Para imágenes a color, se puede aplicar a cada canal por separado, sujeto a a distorsiones en el color o puede cambiarse a otro espacio de color como el HSV y aplicarlo al canal de V.

2.4 Relación de histogramas

La relación de histogramas es la generalización del caso uniforme mostrado en la sección anterior. Se busca que la pdf de una imagen siga la forma de otra, en la sección anterior la pdf objetivo era una distribución uniforme, en este caso seguirá la pdf de otra imagen. Para esto se sigue utilizando la definición 8 y el algoritmo 1.

Algorithm 1: Algoritmo para relación de histogramas

$I_f \leftarrow$ imagen 1 (fuente);
$I_g \leftarrow$ imagen 2 (objetivo);
$F \leftarrow cdf(I_f)$;
$G \leftarrow cdf(I_g)$;
$L =$ niveles de grises;
$x \leftarrow 0$;
$y = [0, L-1]$;
$M \leftarrow$ arreglo de ceros de tamaño L (mapeo);
while $x \leq L-1$ **do**
  Encontrar y tal que $F(x) = G(y)$;
  $M(x) = y$
end
$pixel \leftarrow 1$;
$N \leftarrow$ número de pixeles en I_f ;
I_m imagen mapeada;
while $pixel \leq N$ **do**
  $I_m(pixel) = M(pixel)$;
end

3 Filtrado en el espacio

El filtrado busca eliminar algunas características y mantener otras. En imágenes se pueden clasificar los filtros en suavizado y afilado. Los filtros de suavizado están relacionados con las frecuencias bajas de la imagen y los de afilado con las altas. La operación para aplicar un filtro en el espacio es la convolución la cual definimos en seguida.

3.1 Convolución

La convolución bidimensional es una operación binaria que es el resultado de tres operaciones . Es especialmente útil en el análisis de imágenes, filtrado y detección de características.

La definición formal de la convolución bidimensional discreta es la siguiente:

$$(f * g)(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m, n) g(x - m, y - n)$$

donde $f(x, y)$ y $g(x, y)$ son funciones. En el filtrado espacial una de ellas la imagen y la otra filtro o kernel. Este proceso puede interpretarse como deslizar el kernel $g(x, y)$ sobre la función $f(x, y)$, calculando en cada posición la suma ponderada de los valores de f y de g.

Propiedades importantes:

- Conmutatividad: $f * g = g * f$
- Asociatividad: $(f * g) * h = f * (g * h)$
- Distributividad: $f * (g + h) = f * g + f * h$
- Desplazamiento: $f(x - a, y - b) * g(x, y) = (f * g)(x - a, y - b)$