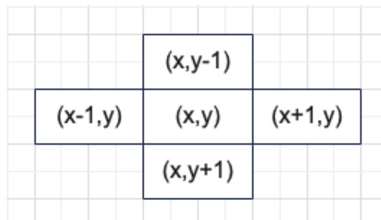
2.3. Transformaciones geométricas de imágenes digitales

Para hablar de transformaciones geométricas de imágenes digitales debemos hablas de pixeles vecinos

Pixeles vecinos

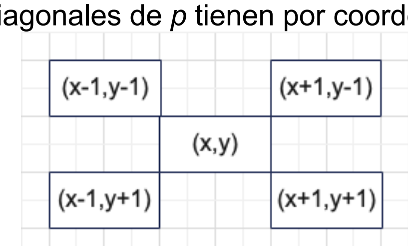
Un píxel *p* obtenido en las coordenadas (x, y) de una imagen *f*(*x*, *y*) tiene cuatro pixeles vecinos, dos pixeles horizontales y dos pixeles verticales cuyas coordenadas vienen dadas por



$$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$$

Este conjunto de píxeles, se denomina los 4 pixeles vecinos de *p*, se denota por $N_4(p)$. Cada píxel tiene una unidad de píxel de distancia del centro (x, y), y algunas de las ubicaciones vecinas de *p* se encuentran fuera la imagen digital si está en los límites superior, inferior, lateral izquierdo y lateral derecho de la imagen.

Los cuatro pixeles vecinos diagonales de *p* tienen por coordenadas a



$$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

y se denotan por $N_D(p)$. Estos puntos, junto con los 4-vecinos, son nombrados como los 8-vecinos de *p*, denotados por $N_8(p)$. Como se mencionó anteriormente, para estos pixeles vecinos algunas de las ubicaciones pueden localizarse dentro y fuera de la imagen si está en los límites de esta.

Adyacencia, conectividad, regiones y límites

Sea *V* el conjunto de valores de intensidad del pixel utilizados para definir la adyacencia. En una imagen binaria, $V = \{1\}$ en referencia a la adyacencia de píxeles con valor 1. Por otra parte, si se trata de una imagen en escala de grises, la idea es la misma, pero el conjunto *V* normalmente contiene más elementos. Por ejemplo, en la adyacencia de píxeles será con un rango de valores de intensidad en el rango de 0 a 255, se establece *V* como un subconjunto de estos 256 valores. Considere tres tipos de adyacencia:

a) 4-adyacencia. Dos píxeles *p* y *q* con valores de *V* son 4-adyacentes si *q* está en el conjunto $N_4(p)$

b) 8-adyacencia. Dos píxeles *p* y *q* con valores de *V* son 8-adyacentes si *q* está en el conjunto $N_8(p)$

c) m-adyacencia (adyacencia mixta). Dos píxeles *p* y *q* con valores de *V* son m-adyacente si:

i) *q* está en el conjunto $N_4(p)$

ii) *q* está dentro de $N_D(p)$ y el conjunto $N_4(p) \cap N_4(q)$ no tiene píxeles cuyos valores son de *V*

La adyacencia mixta es una modificación de la adyacencia en 8, es introducida para eliminar las ambigüedades que surgen cuando se utiliza la adyacencia en 8.

Medida de distancias

Dados los píxeles *p*, *q* y *z*, con coordenadas (*x*, *y*), (*s*, *t*) y (*v*, *w*), respectivamente, y *D* como una función de distancia o métrica si

a) $D(p,q) \geq 0 \quad (D(p,q) = 0 \quad si\, p = q\,)$

b) $D(p,q) = D(q,p)$ y

c) $D(p,z) \leq D(p,q) + D(q,z)$

La distancia euclidiana entre los pixeles *p* y *q* queda definida por

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2} \tag{2.29}$$

En esta medición de distancia, los píxeles cuya distancia es menor o igual a algún valor *r* de (*x*, *y*) se representan en un disco de radio *r* en (*x*, *y*).

La distancia D_4 (es la distancia de manzana de ciudad o distancia de bloque de ciudad) entre los pixeles *p* y *q* se define como

$$D_4(p,q) = |x-s| + |y-t| \tag{2.30}$$

Los píxeles que tienen una distancia D_4 desde (*x*, *y*) menor o igual a algún valor *r* y forman un rombo centrado en (*x*, *y*). Los píxeles con distancia $D_4 \leq 2$ desde (*x*, *y*) (el punto central) forman los siguientes contornos de distancia constante:

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

Los píxeles con $D_4 = 1$ son los 4 vecinos de (*x*, *y*).

La distancia D_8 (nombrada distancia de tablero de ajedrez) entre los pixeles *p* y *q* se define como

$$D_8(p,q) = max(|x-s|, |y-t|) \tag{2.31}$$

Los píxeles con distancia D_8 desde (*x*, *y*) menor o igual a algún valor *r* forman un cuadrado con centro en (x,y). Por ejemplo, los píxeles con distancia $D_8 \leq 2$ desde (*x*, *y*) (el punto central) forman los siguientes contornos de distancia constante:

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

Los píxeles con $D_8 = 1$ son los ocho vecinos de (x, y).

Las transformaciones geométricas cambian la distancia entre los píxeles de una imagen. Estos cambios se denominan cambios de papel elástico y son similares a imprimir una imagen en papel de goma y estirar el papel según las reglas comentadas anteriormente. En el contexto del procesamiento de imágenes digitales, las transformaciones geométricas tienen dos funciones principales:

1) Una transformación espacial de coordenadas.
2) Interpolación de intensidad que asigna valores de intensidad a los píxeles transformados espacialmente.

La transformación de coordenadas se puede expresar de la siguiente forma

$$(x, y) = T\{(v, w)\} \tag{2.32}$$

donde (v, w) son coordenadas de píxeles en la imagen original y (x, y) son las correspondientes coordenadas de los píxeles en la imagen transformada. por ejemplo, la transformación $(x, y) = T\{(v, w)\} = (v/2, w/2)$ reduce la imagen original a la mitad de su tamaño en ambas direcciones espaciales. Uno de los espacios más utilizados. transformaciones de coordenadas es la transformada afín (de Wolberg), que tiene la forma general

$$[x \;\; y \;\; 1] = [v \;\; w \;\; 1]T = [v \; w \; 1]\begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix} \tag{2.33}$$

Esta transformación puede escalar, rotar, trasladar o distorsionar un conjunto de coordenadas de
puntos, dependiendo del valor elegido para los elementos de la matriz **T**.

2.3.1. Interpolación

La interpolación es una herramienta básica utilizada en tareas como hacer zoom, reducir, corrección en rotación y corrección geométrica. El objetivo principal en esta sección es introducir la interpolación y aplicarla al cambio de tamaño de la imagen (reducción y zoom), que son básicamente métodos de re muestreo de una imagen digital. La interpolación es el proceso de usar datos conocidos para estimar valores en lugares desconocidos. Suponga que una imagen de tamaño 500 X 500 pixeles tiene que ser ampliada 1,5 veces, es decir, a 750 X 750 píxeles. Una forma sencilla de visualizar el zoom es crear una cuadrícula imaginaria con el mismo espaciado de 750 X 750 píxeles que el original, y luego encójalo para que encaje exactamente sobre la imagen original. En el espaciado de los píxeles en la cuadrícula reducida será menor que el espaciado de píxeles en la imagen original. Para asignar la intensidad de cada pico, busque el píxel más cercano en la imagen original y asigne la intensidad de ese píxel a un nuevo píxel en la cuadrícula. Cuando se haya terminado el asignando de intensidades a todos los puntos en la cuadrícula superpuesta, se expande al tamaño original especificado para obtener la imagen ampliada.

El algoritmo descrito anteriormente se llama definición local porque asigna nuevas regiones al tamaño de los vecinos del píxel de la imagen original. Este método tiende a producir píxeles innecesarios, como una curvatura excesiva de líneas rectas, y por lo tanto no se utiliza. El mejor método es la interpolación bidireccional, donde se utilizan cuatro vecinos para estimar la intensidad de un punto. Denote (x, y) a las coordenadas de la ubicación que se quiere asignar un valor de intensidad (piense en ello como un punto de la cuadrícula descrita anteriormente), y $v(x, y)$ denotar ese valor de intensidad, para la interpolación bilineal, el valor asignado se obtiene mediante la ecuación

$$v(x, y) = ax + by + cxy + d \qquad (2.34)$$

donde los cuatro coeficientes se determinan a partir de las cuatro ecuaciones con cuatro incógnitas, se pueden describir usando los cuatro pixeles vecinos más cercanos del punto (x, y). La interpolación bilineal da mejores resultados que el vecino más cercano, o interpolación vecina, con un aumento modesto en la carga computacional. El siguiente nivel de complejidad es la interpolación bicúbica, que involucra los dieciséis vecinos más cercanos de un punto (x, y). El valor de intensidad asignado al punto es obtenido usando la ecuación

$$v(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^i \qquad (2.35)$$

Dieciséis de los coeficientes están determinados por dieciséis ecuaciones y se pueden escribir dieciséis incógnitas utilizando los vecinos más cercanos del punto (x, y). Observe que la ecuación (2.30) se reduce en forma a la ecuación (2.29) si los límites de ambas sumas en la ecuación anterior son 0 a 1. Generalmente, la interpolación bicúbica hace un mejor trabajo preservando los detalles finos que su bilineal contrapartida. La interpolación bicúbica es el estándar utilizado en la imagen comercial. programas de edición, como Adobe Photoshop y Corel Photopaint.

2.3.2. Traslación.

Empleando la ecuación 2.32 la matriz empleada es

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad \text{con ecuaciones} \begin{bmatrix} x = v + t_x \\ y = w + t_y \end{bmatrix} \qquad (2.36)$$

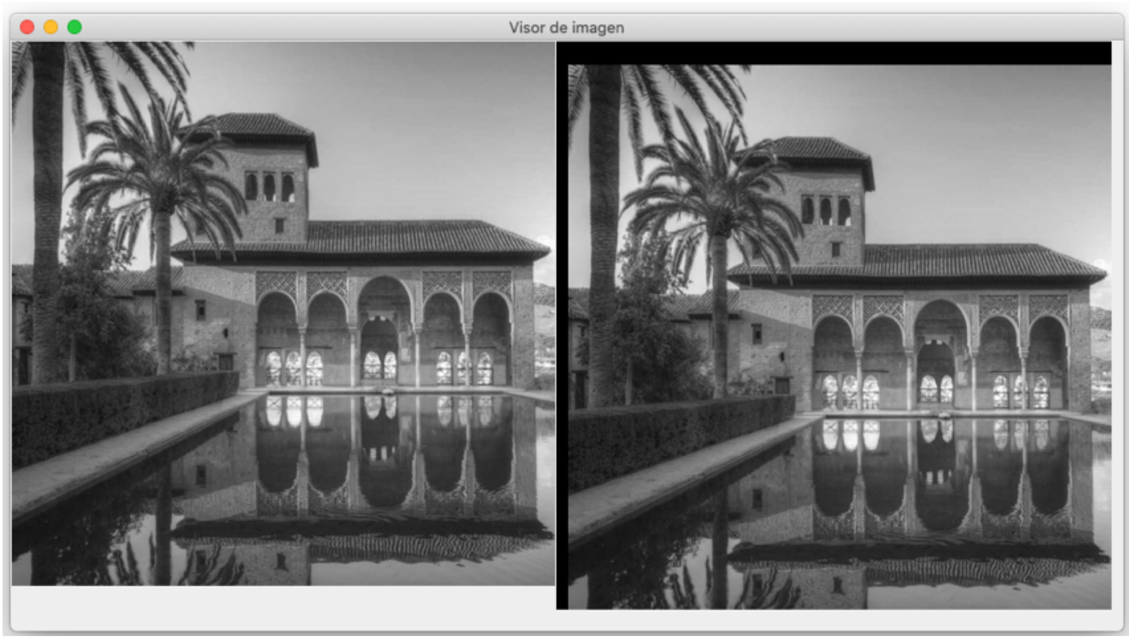


Figura 2.19 Traslación de la imagen con $t_x = 10$ y $t_y = 20$ pixeles respectivamente.

2.3.3. Rotación

Empleando la ecuación 2.32 la matriz empleada es

$$\begin{bmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{con ecuaciones} \begin{bmatrix} x = v\ cos\theta - w\ sin\theta \\ y = v\ cos\theta + w\ sin\theta \end{bmatrix} \tag{2.37}$$

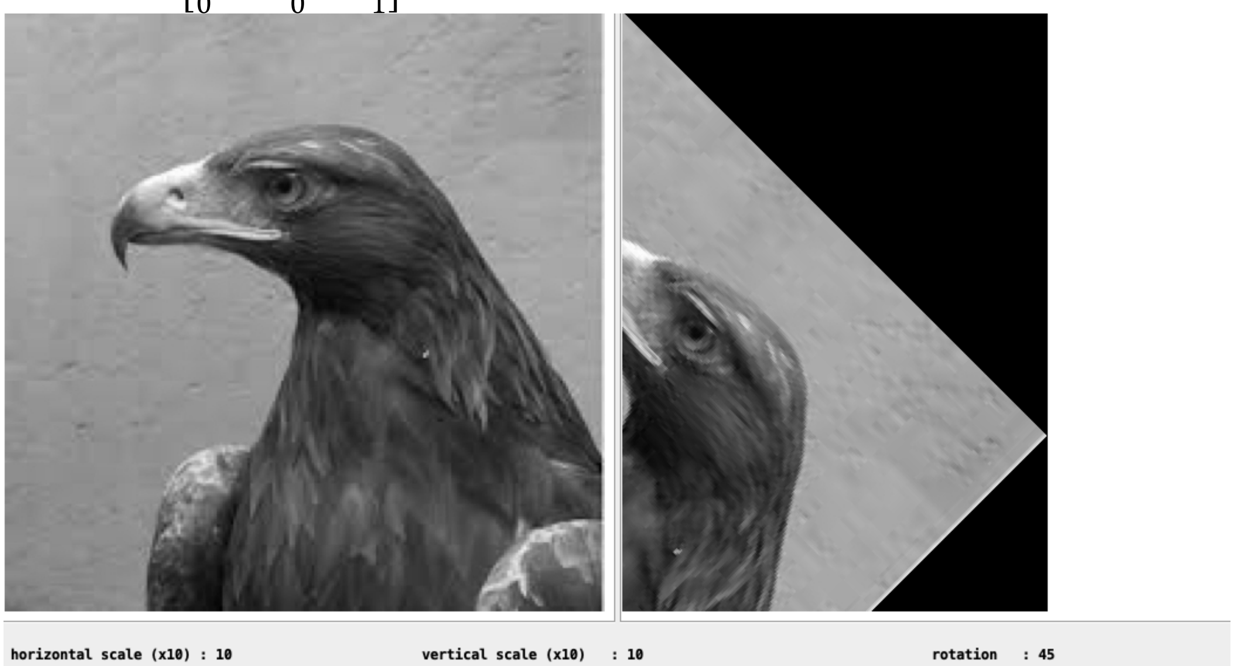


Figura 2.20 Rotación de la imagen respecto al origen 45 grados en sentido horario.

Figura 2.21 Rotación de la imagen respecto al origen 45 grados en sentido anti horario.

2.3.4. Escalamiento

Empleando la ecuación 2.32 la matriz obtenida es

$$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad \text{con ecuaciones} \begin{bmatrix} x = c_x v \\ y = c_y w \end{bmatrix} \tag{2.38}$$

Figura 2.22 Escalado con cx=1.5 y cy=1.5 de la imagen original con interpolación de orden cero.

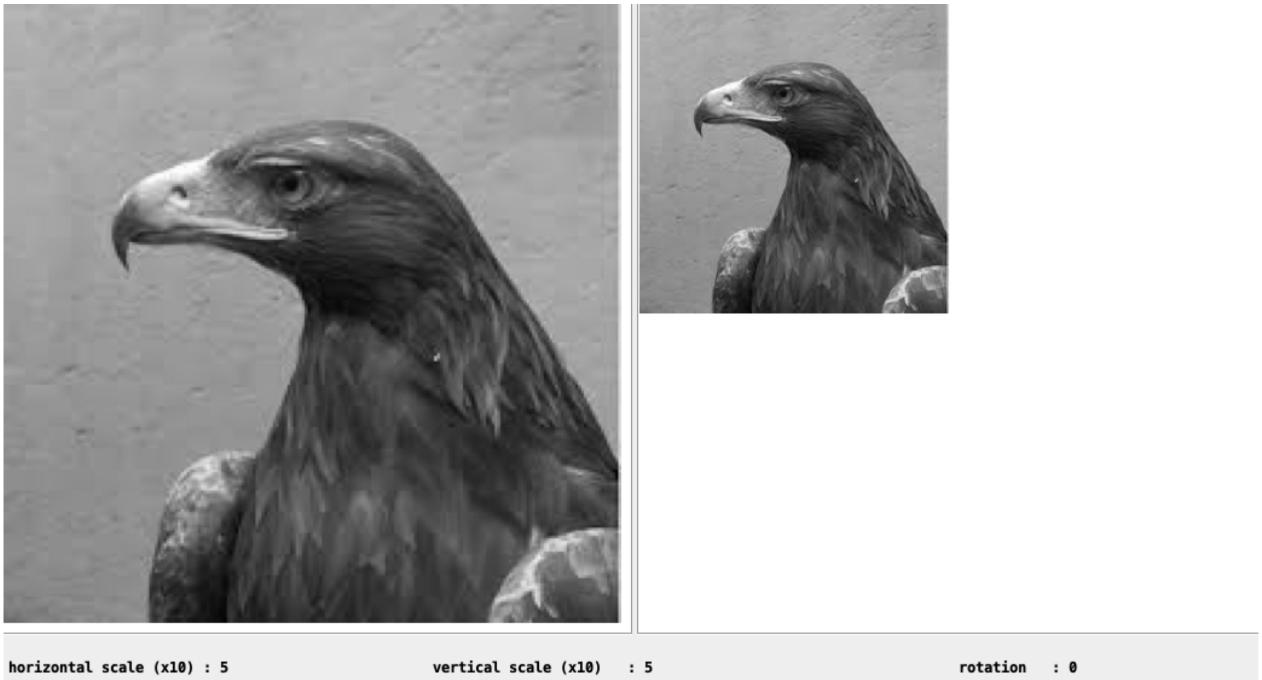


Figura 2.23 Escalado con cx=0.25 y cy=0.25 de la imagen original con interpolación de orden cero.

Figura 2.24 Escalado con cx=1.5 y cy=1.5 de la imagen original con interpolación de primer orden.

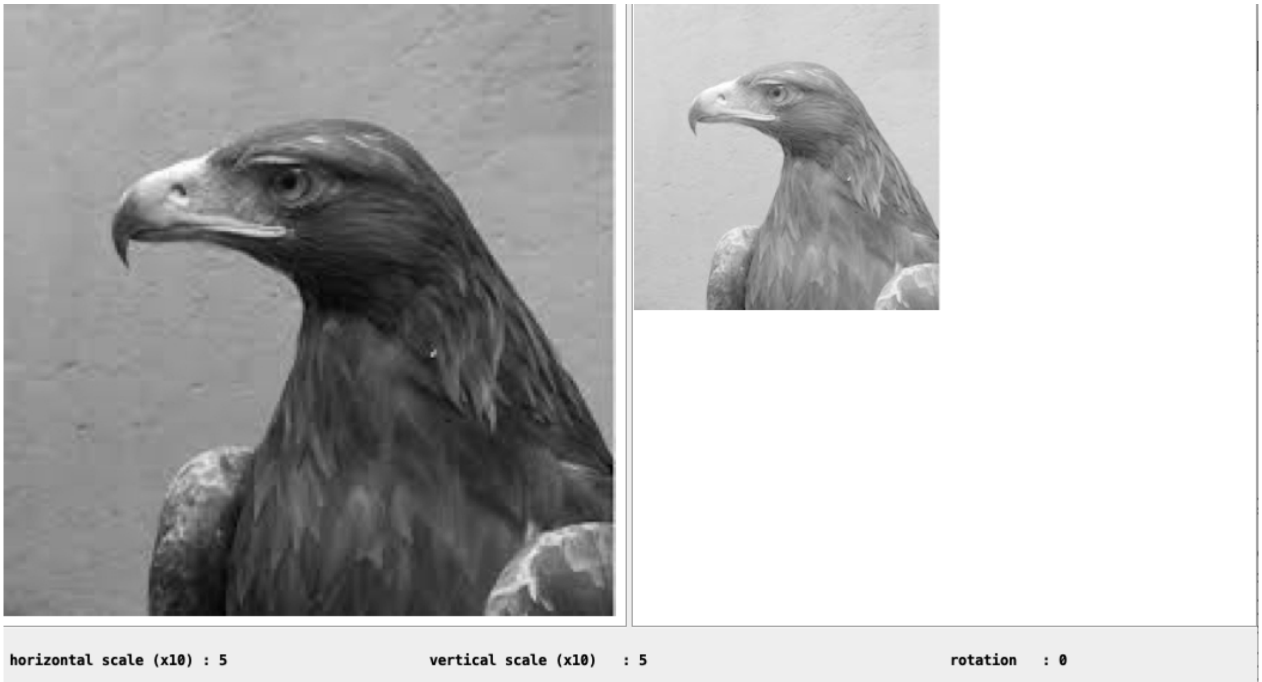


Figura 2.25 Escalado con cx=0.25 y cy=0.25 de la imagen original con interpolación de primer orden.