
Universidad
Tecmilenio®

Universidad Tecmilenio

Visión computacional

Introducción al procesamiento de imágenes

Introducción

El propósito principal de la **visión computacional** es obtener información de los pixeles existentes en una imagen digitalizada.

Desde la perspectiva en ingeniería, permite construir sistemas autónomos que puedan realizar algunas de las tareas que ejecuta el sistema visual humano.

En este tema conocerás:
• Conceptos importantes en el procesamiento de imágenes.
• Formato de representación de imágenes.

Explicación

Para los humanos, la obtención de información de una imagen es un proceso trivial. Cualquier persona puede describir con facilidad lo que ve en una fotografía, reconocer un rostro aun cuando solo lo ha visto una vez, incluso los niños pueden platicar de una película al haberla visto también una vez. Para realizar estas tareas, los seres humanos utilizan la retina, el nervio óptico y el córtex visual. Con la finalidad de replicar este comportamiento, en la visión computacional se cuentan con cámaras, datos y algoritmos que son capaces de obtener la información de una imagen.

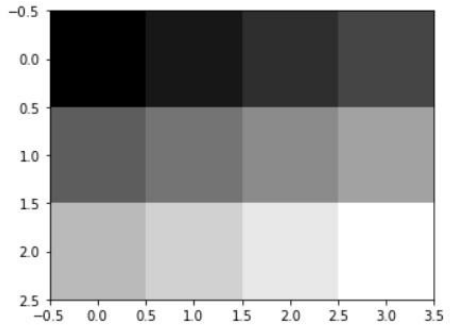
Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

La representación de imágenes se hace por medio de coordenadas. En matemáticas se tiene la costumbre que el eje **x** va de izquierda a derecha y el eje **y** de abajo hacia arriba con el origen ubicado en la esquina inferior izquierda, sin embargo, para una imagen digital es distinto.

En la figura se utilizan las librerías numpy y matplotlib para construir la imagen representada por las muestras definidas en la matriz siguiente.

```
In [2]: import numpy as np

        f = np.arange(12).reshape(3,4)

In [3]: f

Out[3]: array([[ 0,  1,  2,  3],
               [ 4,  5,  6,  7],
               [ 8,  9, 10, 11]])

In [6]: import matplotlib.pyplot as plt

        plt.imshow(f, cmap='gray', interpolation='nearest');
        plt.show()
```

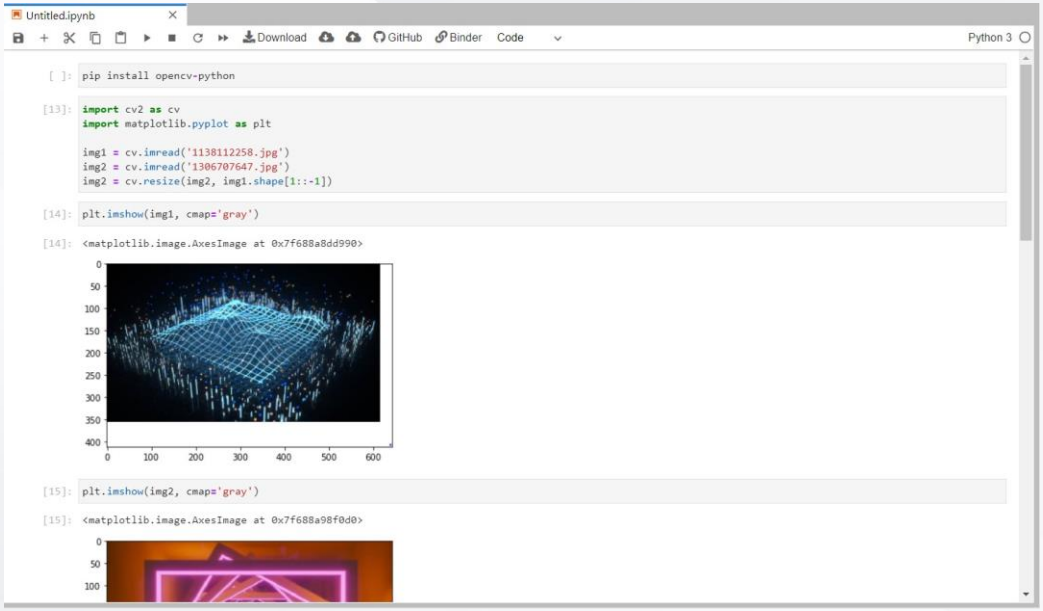



```
In [ ]:
```

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

Los **operadores de puntos** permiten realizar operaciones matemáticas para modificar imágenes pixel a pixel (el punto). Estas operaciones son básicamente sumas y restas ampliamente utilizadas en procesamiento de imágenes: mezclado alfa (*α-Blending*),



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

Enmascaramiento de enfoque (*unsharp masking)*

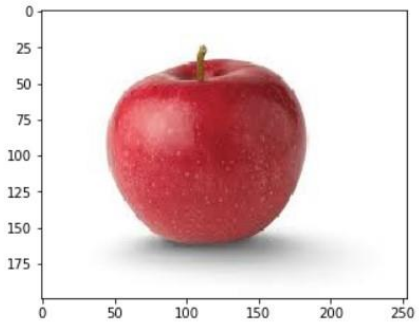
Untitled.ipynb

Download GitHub Binder Code Python 3

```
img1 = cv.imread('manzana.jpg')
```

[79]:
```
plt.imshow(cv.cvtColor(img1, cv.COLOR_BGR2RGB), cmap='gray')
```

[79]: <matplotlib.image.AxesImage at 0x7f6889a8b750>



[81]:
```
gauss = cv.GaussianBlur(img1, (7,7), 0)
unsharp_img1 = cv.addWeighted(img1, 2, gauss, -1, 0)
```

[82]:
```
plt.imshow(cv.cvtColor(unsharp_img1, cv.COLOR_BGR2RGB), cmap='gray')
```

[82]: <matplotlib.image.AxesImage at 0x7f6889a57e50>

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

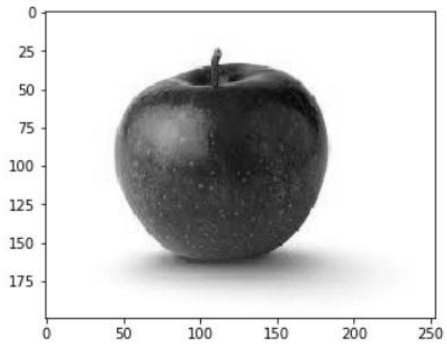
Universidad Tecmilenio

Umbralización de imágenes (*thresholding*).

Untitled.ipynb

Download GitHub Binder Code Python 3

```
[7]: import cv2 as cv
     import numpy as np
     import matplotlib.pyplot as plt

[8]: img1 = cv.imread('manzana.jpg')

[9]: img = cv.cvtColor(img1, cv.COLOR_BGR2GRAY)
     plt.imshow(img, cmap='gray')

[9]: <matplotlib.image.AxesImage at 0x7fbfb00748d0>
```



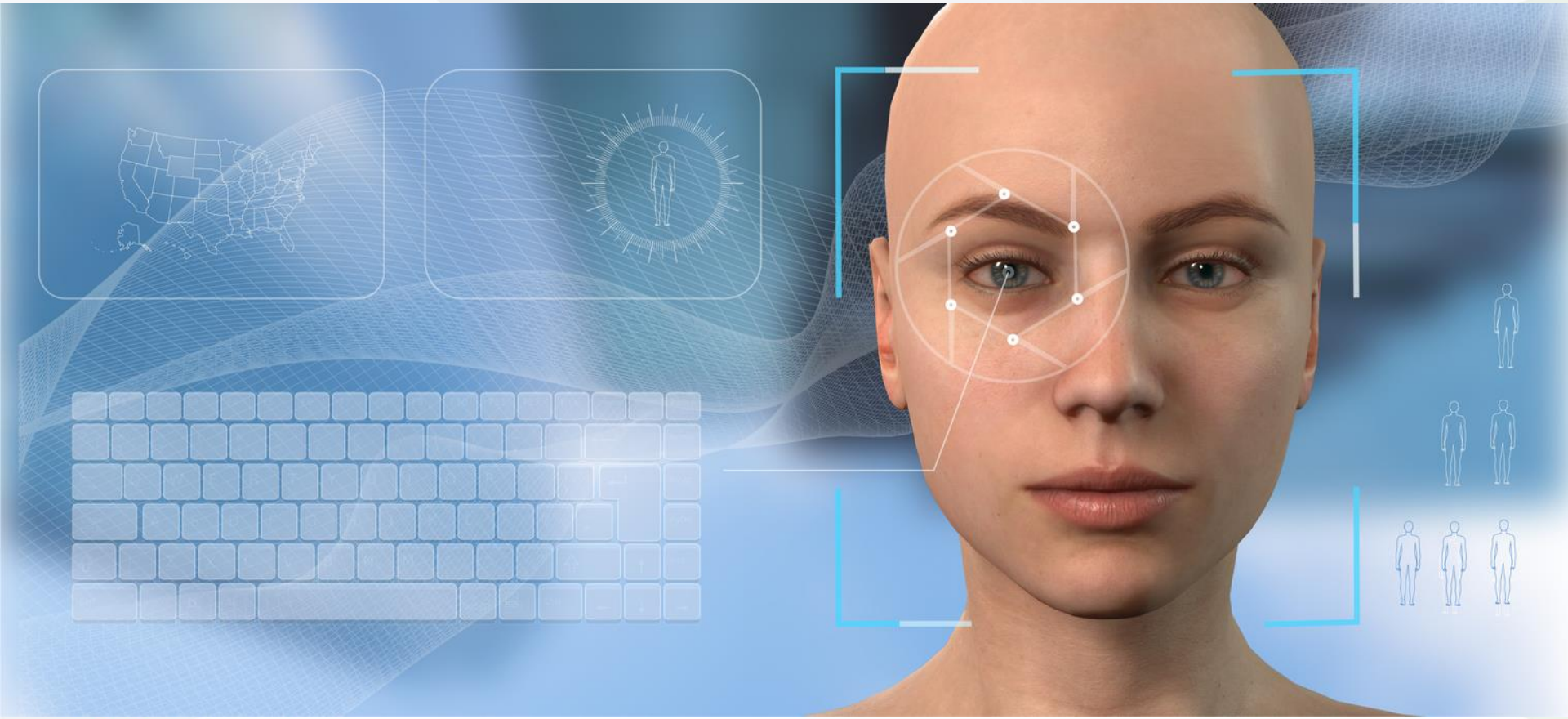

```
[10]: ret, thresh1 = cv.threshold(img, 120, 255, cv.THRESH_BINARY)
      ret, thresh2 = cv.threshold(img, 120, 255, cv.THRESH_BINARY_INV)
      ret, thresh3 = cv.threshold(img, 120, 255, cv.THRESH_TRUNC)
      ret, thresh4 = cv.threshold(img, 120, 255, cv.THRESH_TOZERO)
      ret, thresh5 = cv.threshold(img, 120, 255, cv.THRESH_TOZERO_INV)
```

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Actividad

¿Sabes qué es un espacio de color?¿Cuántos conoces y cuántos existen?

Averigua cómo funciona (matemáticamente) la función cv.COLOR_BGR2GRAY



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Cierre

El procesamiento de imágenes es un campo que resuelve tareas o problemas con las imágenes, e incluye el tratamiento a distintos tipos de imagen, discretización, interpolación y coloración. Además, es un campo que procesa imágenes a partir de transformaciones, las mejora y ubica puntos de interés en ellas.

Recuerda que una imagen digital puede ser representada como un arreglo y que, a su vez, puede contener muestras de imágenes.

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Universidad Tecmilenio

Visión computacional

Operadores geométricos

Introducción

Las transformaciones geométricas son de las operaciones más comunes que se encuentran en cualquier tarea de procesamiento de imágenes; modifican las imágenes manteniendo los colores, pero colocando los pixeles en lugares distintos.

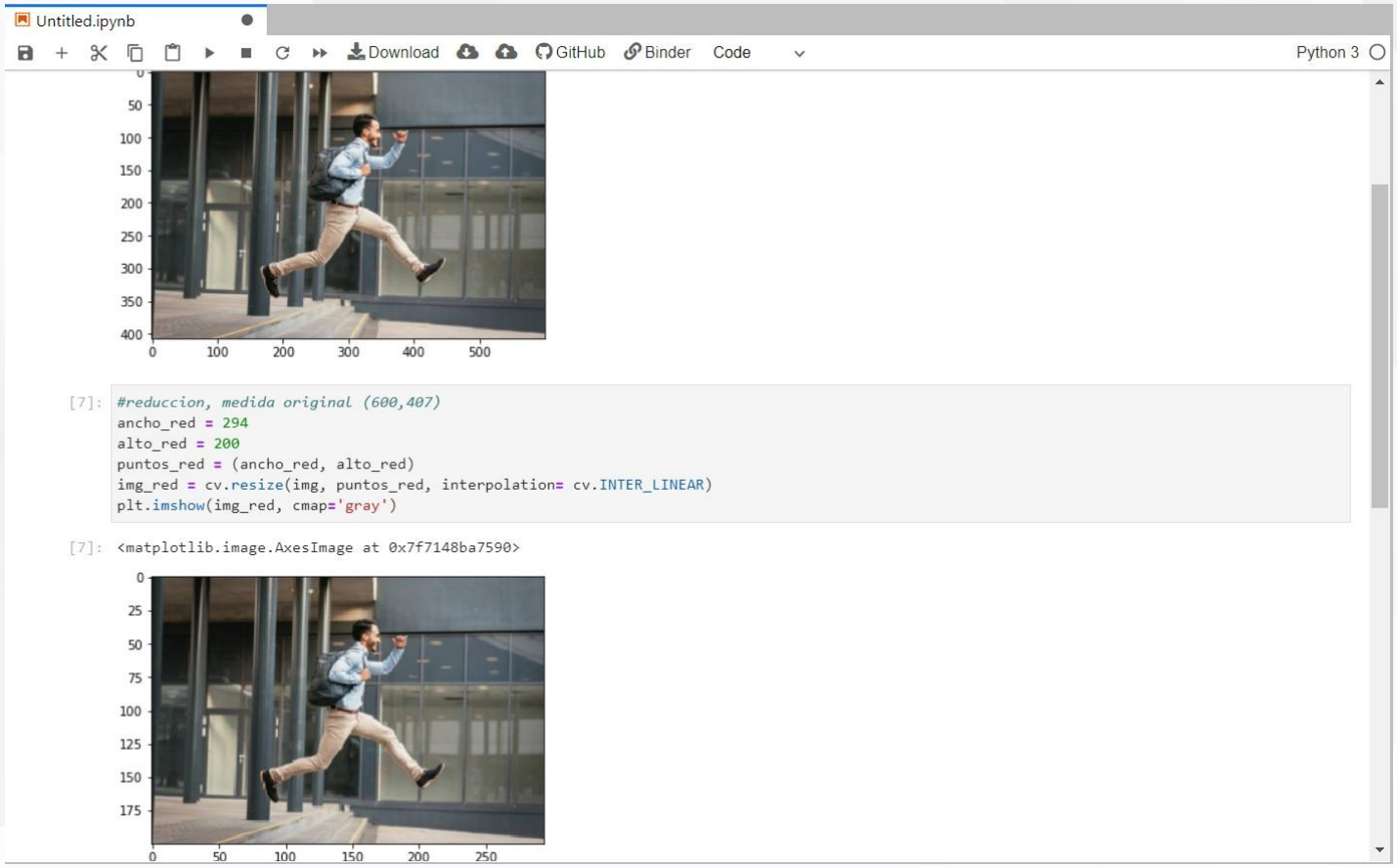
En este tema conocerás:
- Los diferentes tipos de operadores geométricos.
- La representación en coordenadas homogéneas de los operadores geométricos.
- El uso de OpenCV en Python para realizar transformaciones u operaciones geométricas.

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

Los operadores geométricos son rotación, traslación y escalamiento, y se conocen como transformaciones geométricas afines.

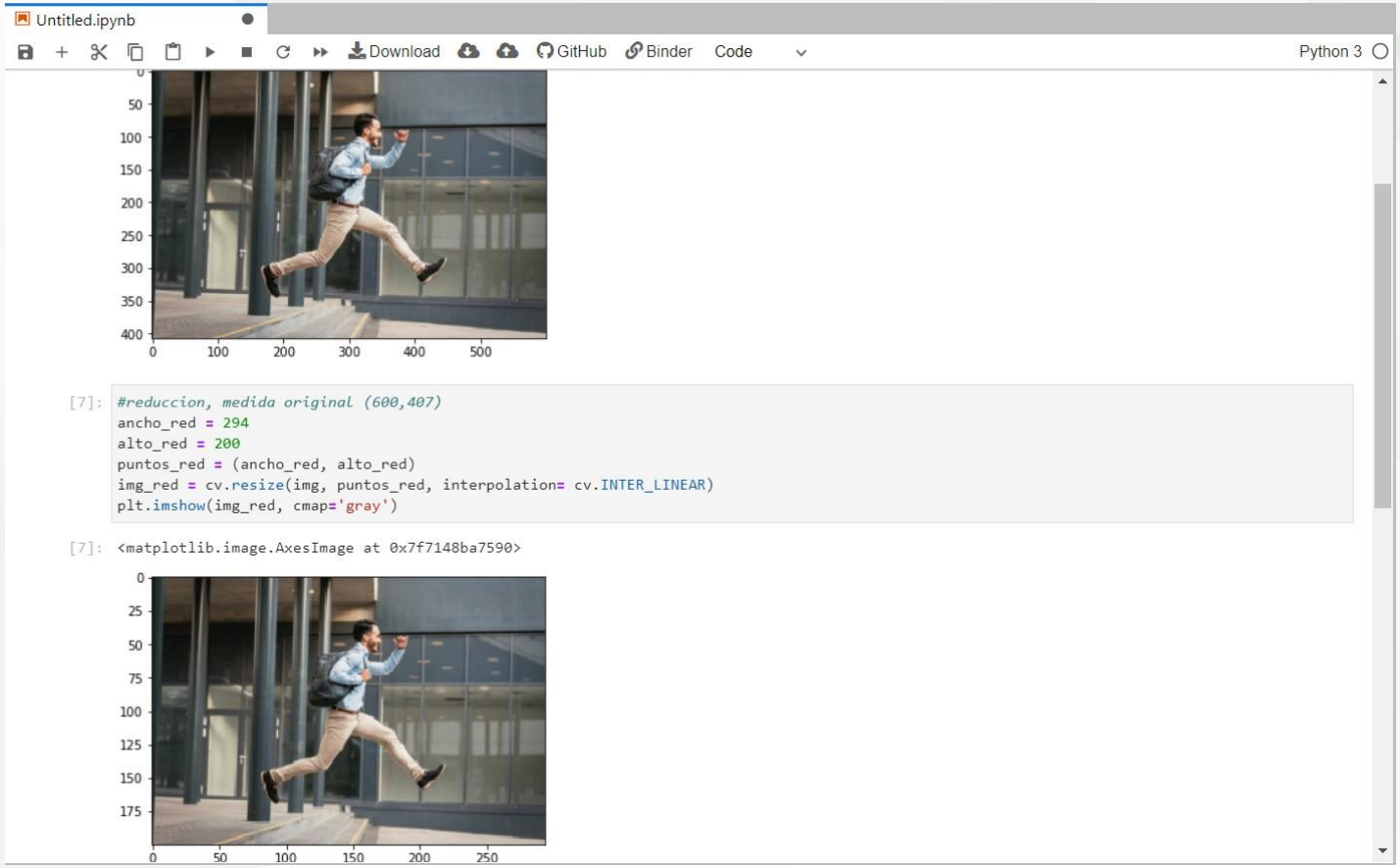
Desde el punto de vista geométrico, la **rotación** es el movimiento circular de un objeto alrededor de un punto y se caracteriza por el **ángulo de rotación** (la amplitud de rotación), el **centro de rotación** (un punto de referencia) y el **sentido de la rotación**. El plano geométrico a lo largo del cual ocurre la rotación se conoce como **plano de rotación** y la línea imaginaria perpendicular a este, **eje de rotación**.



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

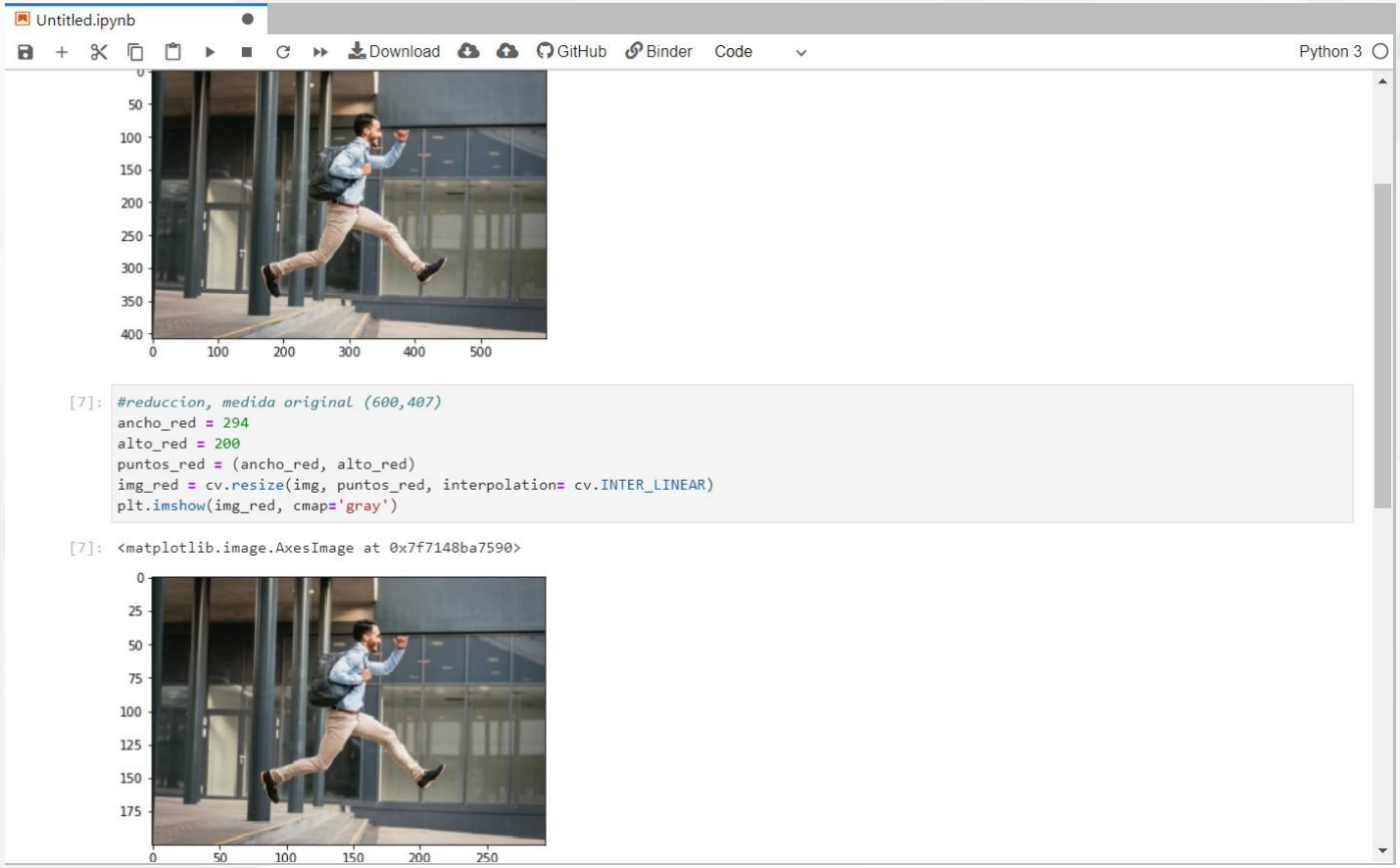
La **traslación** es un movimiento de desplazamiento. Geométricamente, es una transformación que desplaza los puntos de una figura a una cierta distancia en una dirección específica.
En visión computacional, la traslación de una imagen implica desplazarla una cierta cantidad de pixeles a lo largo de alguno de los ejes *x, y* del plano cartesiano.



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

La transformación de **escalamiento** aumenta o disminuye la extensión espacial de la imagen. Para modificar el tamaño de la imagen se requiere escalarla a lo largo de cada uno de sus ejes (alto y ancho), según el factor de escala especificado. Si, por ejemplo, la imagen fuera escalada por un factor de 2, entonces usaría el doble de pixeles en cada dimensión para contener la información.



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Explicación

Las coordenadas homogéneas permiten describir un punto en el espacio y convierten el punto cartesiano (x,y) a una representación homogénea, utilizando tres parámetros (Juárez y Díaz, 2017). Estas coordenadas facilitan la representación de las ecuaciones utilizadas en las transformaciones geométricas, usando multiplicaciones de vectores y matrices.

la matriz de rotación

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

la matriz de traslación

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & T_x \\ 0 & 1 & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

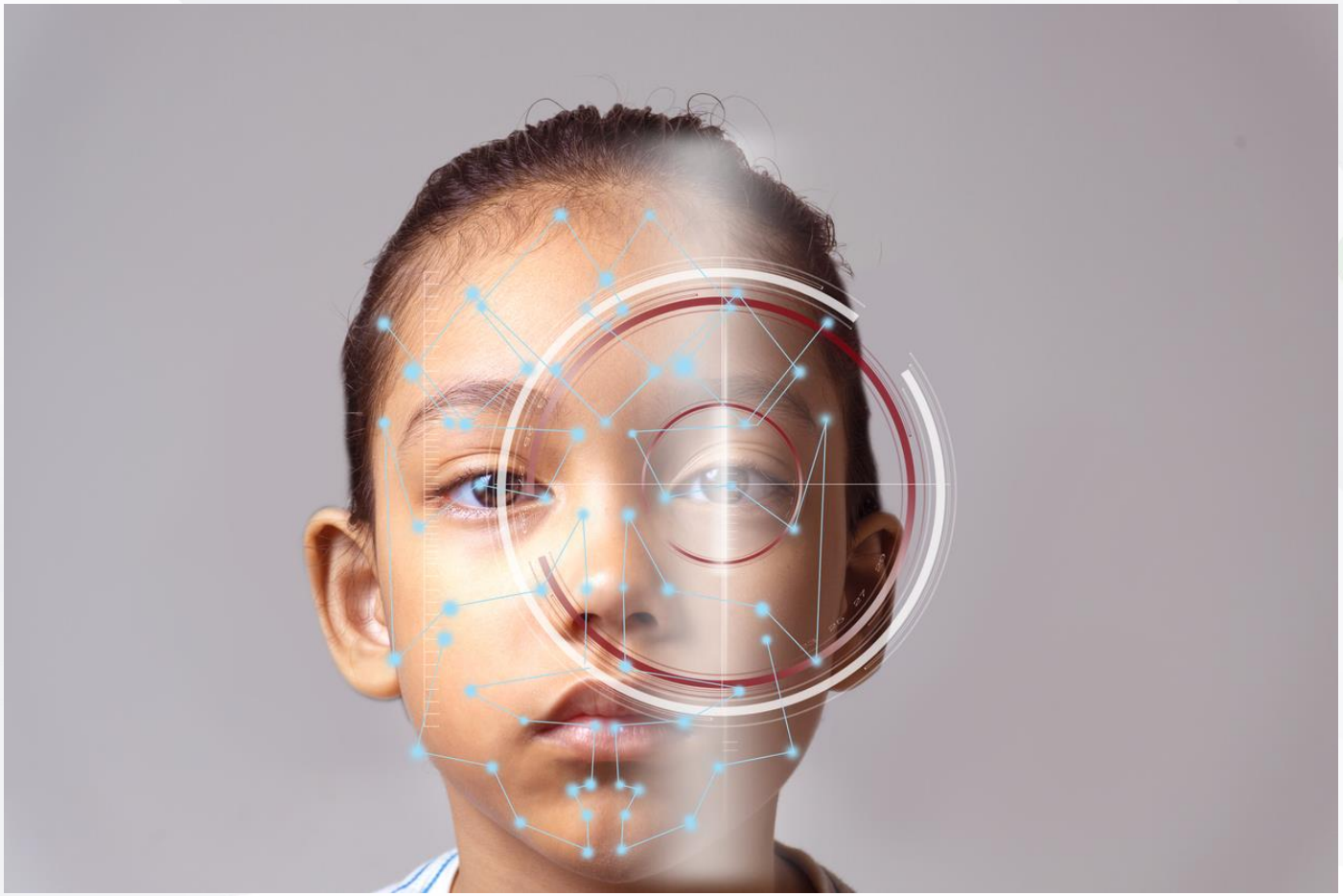
la matriz de escalamiento

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Actividad

Cuando se realizó la rotación y la traslación, la imagen resultante apareció recortada, por lo tanto, desarrolla e implementa un algoritmo para que las imágenes se puedan visualizar de manera completa.



Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .

Cierre

Los operadores geométricos mueven los pixeles a ubicaciones nuevas dentro de la imagen en lugar de alterar la intensidad de cada pixel.
Los operadores más comunes se definen a partir de transformaciones afines y con ellas es posible modificar las dimensiones de una imagen, girarla, desplazarla o inclinarla. Las matrices proporcionan un marco formal para trabajar con problemas geométricos en visión computacional.

Derechos de Autor Reservados. (UNIVERSIDAD TECMILENIO®) .