Image Transformation

Tushar B. Kute,
http://tusharkute.com

Image Translation

- Image translation in computer vision refers to the process of transforming an image into another image while preserving certain visual characteristics or changing them in a controlled manner.
- This transformation can involve various manipulations

Image Translation

- Style transfer: Applying the artistic style of one image to another (e.g., making a photo look like a watercolor painting).
- Domain adaptation: Adapting images from one domain (e.g., satellite images) to another domain (e.g., street view images).
- Super-resolution: Generating a higher-resolution image from a lower-resolution one.
- Inpainting: Filling in missing or damaged parts of an image.
- Image colorization: Adding color to a grayscale image.

Image Translation: Technologies

- Deep learning:
  - Deep neural networks, particularly Generative Adversarial Networks (GANs), are widely used for image translation tasks. These networks learn the relationship between source and target images and can generate realistic results.
- Convolutional neural networks (CNNs):
  - CNNs excel at extracting spatial features from images, making them a suitable choice for tasks like style transfer and domain adaptation.
- Autoencoders:
  - These neural networks learn a compressed representation of an image and can be used for tasks like image denoising and super-resolution.

Image Translation: Applications

- Artistic image editing: Users can apply different artistic styles to their photos for creative expression.
- Medical imaging: Image translation can be used to enhance medical images for better diagnosis and treatment planning.
- Autonomous vehicles: Image translation can help self-driving cars adapt to different weather conditions and environments.
- Video game development: Image translation can be used to create game assets with different styles and textures.
- Image restoration: Damaged or old photos can be repaired and restored using image translation techniques.

Translation Matrix

- In OpenCV, a translation matrix is a 2x3 matrix used to perform various kinds of image transformations, including translating an image, applying affine transformations, and adjusting perspective.
- It plays a crucial role in manipulating images and achieving desired visual effects.

Translation Matrix

- Here's a breakdown of its elements:
  - 2 rows: Represent the transformation applied to the horizontal (x) and vertical (y) dimensions of the image.
  - 3 columns: Represent the original coordinates of a point in the image (x, y, 1) and the resulting transformed coordinates after applying the matrix.

Translation Matrix

- The specific values within the matrix determine the nature of the transformation:
- First row:
  - [1, 0, Tx]: Shifts the image Tx pixels to the right while preserving its height.
  - [0, 1, Ty]: Shifts the image Ty pixels down while preserving its width.
- Second row:
  - A combination of values in both columns can achieve more complex translations and tilts.

Translation Matrix

- Let's see an example of translating an image by 50 pixels to the right:
- translation_matrix = np.float32([[1, 0, 50], [0, 1, 0]])
- Here, the [1, 0, 50] row indicates no scaling or shearing in the horizontal direction but a shift of 50 pixels to the right. The second row [0, 1, 0] signifies no change in the vertical dimension.
- OpenCV provides functions like cv2.warpAffine and cv2.warpPerspective that take this translation matrix as an input and apply the corresponding transformation to the image.

Translation

- Practical

Image Rotation

- Image rotation refers to the process of turning an image around a fixed point, resulting in a new image with the original information rearranged in a different orientation.
- This is a common operation in image processing used for various purposes.

Image Rotation: Why?

- Correcting camera tilt: Rotating an image to compensate for the camera being tilted during capture results in a more natural-looking perspective.
- Alignment and organization: Images can be rotated to align features of interest or organize them in a specific layout.
- Data augmentation: Image rotation can be used as part of data augmentation techniques to increase the diversity of training data for machine learning models.
- Artistic effects: Rotating images can create interesting visual effects and distortions for artistic expression.

Image Rotation: Types

- There are two main types of image rotation:
  - Rigid rotation: The entire image rotates around a fixed center point, maintaining its original shape and proportions. This is the most common type of rotation.
  - Affine rotation: This involves a combination of rotation and shear, altering the image's shape and proportions.

Image Rotation: How

- Rotation Matrix:
  - A rotation matrix, denoted by R, is a 2x2 matrix that encodes the specific rotation angle and direction. The specific values within the matrix depend on the angle and axis of rotation. Here's a general form:
  - R = | cos(theta) -sin(theta) |

    | sin(theta) cos(theta) |
- where theta is the rotation angle in radians.

Image Rotation: How

- Applying the Transformation:
  - To rotate a point (x, y) by applying the rotation matrix R, we perform the following matrix multiplication:

    (rotated_x, rotated_y) = R * (x, y)

  - This multiplication results in a new vector (rotated_x, rotated_y) representing the pixel's location after the rotation.

Image Rotation: How

- Rotating the Entire Image:
  - By applying the rotation matrix to each pixel's vector using the above equation, we effectively rotate the entire image around its center.
  - This process involves iterating through each pixel, performing the multiplication, and updating its coordinates based on the rotated values.

Image Rotation

- Practical

Image Resampling/Rescaling/Resizing

- Image Resizing is the process of increasing or decreasing the size of an image.
- What does it mean?
  - In terms of a 2-dimensional image, more specifically, it means scaling up/down the width and height of an image by computing the pixel values for the newer(or resized) image.

Image Resampling/Rescaling/Resizing

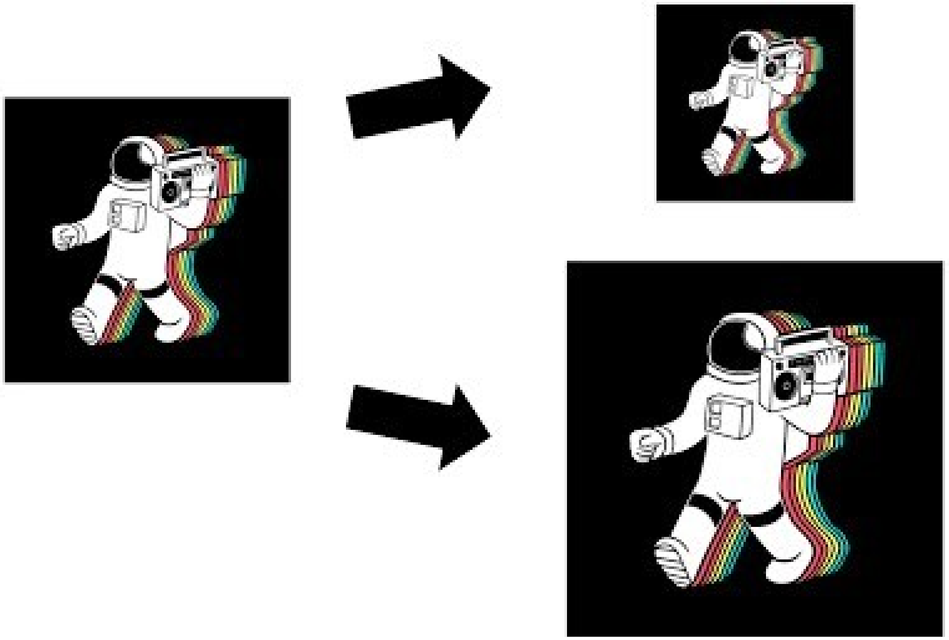

Image Resizing: How?

- Interpolation: This is the primary step in resizing, where missing pixels are estimated based on surrounding values. Popular methods include:
- Nearest neighbor: Simplest method, assigns the value of the nearest existing pixel to the new pixel. Good for preserving sharp edges but can appear blocky.
- Bilinear interpolation: Averages the values of the four nearest pixels, producing smoother results but potentially blurring image details.
- Bicubic interpolation: Uses a weighted average of sixteen surrounding pixels, offering higher quality resizing but computationally more expensive.

Image Resizing: How?

- Transformation matrices: Linear algebra comes into play here. Resizing can be viewed as applying a transformation to the image coordinates. This can be achieved through:
  - Scaling: A scaling matrix multiplies the original coordinates by a factor (e.g., scale by 0.5 for half size), effectively shrinking or enlarging the image.
  - Shearing: Less commonly used, shearing matrices can tilt the image horizontally or vertically.

Image Resizing

- Practical

Image Flipping

- Image flipping in computer vision refers to the process of mirroring an image along a specific axis (horizontal or vertical) or both, resulting in a new image with the original information rearranged around the chosen axis.
- It's a common image manipulation technique with various applications, including:
  - Data augmentation
  - Symmetry Detection
  - Visual effects and artistic expression

Image Flipping

- Types of image flipping:
  - Horizontal flip: Mirrors the image left to right, reversing the positions of objects and features.
  - Vertical flip: Mirrors the image top to bottom, inverting the orientation of objects and features.
  - Diagonal flip: Flips the image along the diagonal axis, creating a mirrored version rotated by 45 degrees.

Image Cropping

- Image cropping in computer vision refers to the process of selecting and extracting a specific region of interest (ROI) from an image and discarding the remaining parts.

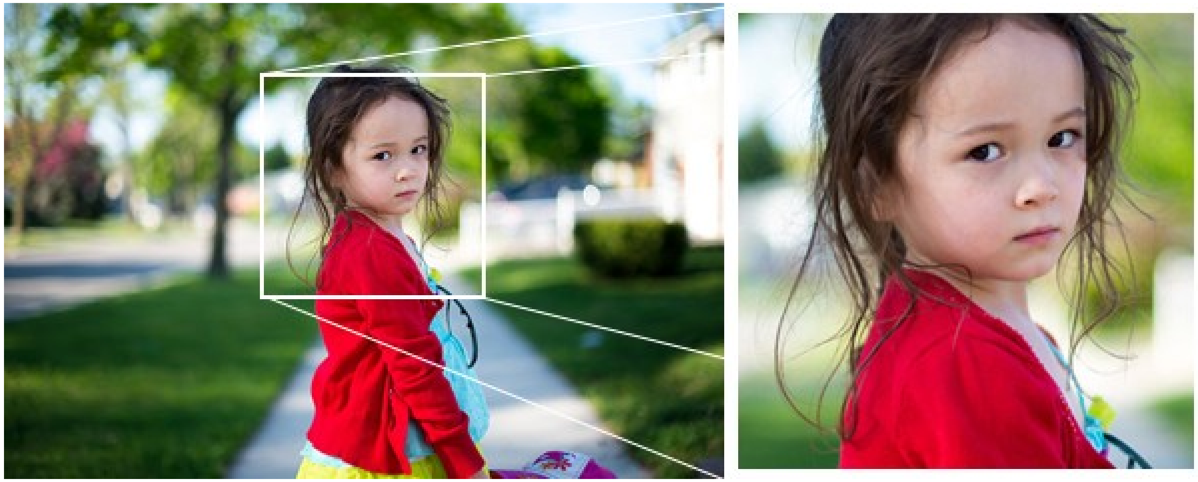

Image Cropping: Why?

- Object detection and recognition: After identifying an object in an image, cropping can focus on the object for better analysis or classification.
- Background removal: Cropping can isolate the subject of interest from distracting backgrounds, improving visual focus and processing.
- Data augmentation: Cropping images from larger ones creates multiple variations for training machine learning models, enhancing their robustness and generalizability.
- Visual composition and editing: Cropping can improve the composition of an image by removing unwanted elements or emphasizing specific features.

Image Cropping: How?

- Bounding boxes: Defining a rectangular box around the desired area is a simple and efficient method. Libraries like OpenCV provide functions for extracting ROI based on bounding box coordinates.
- Polygons: For irregular shapes, defining a polygon with multiple vertices offers more flexibility in selecting the ROI.
- Masking: Creating a mask image where the ROI is white and the rest is black allows precise extraction of the desired area.

Image Cropping: Techniques

- Content-aware cropping: Algorithms can analyze the image content and automatically suggest optimal cropping regions based on saliency or object detection.
- Adaptive cropping: The ROI can be dynamically adjusted based on specific criteria, such as maximizing the presence of objects or minimizing empty space.

Image Cropping

- Practical

Shearing

- Shearing transformation in computer vision is a specific type of geometric transformation that distorts an image by tilting it along a specific axis.
- This creates a slanting effect, as if layers of the image are sliding past each other in one direction.
- Here are some key points about shearing transformation:
- Types of shearing:
  - X-shearing: Tilts the image along the x-axis, causing vertical lines to slant.
  - Y-shearing: Tilts the image along the y-axis, causing horizontal lines to slant.

Shearing

- Effect on objects:
  - Shearing alters the shape and size of objects in the image.
  - It can be used to correct for slight rotations or introduce specific distortions for artistic effects or data augmentation.

Shearing

- Applications:
  - Correcting text slant: Can be used to straighten tilted text in scanned documents.
  - Data augmentation: Creating variations of images by introducing controlled shearing for training machine learning models.
  - Special effects: Used in image editing software to create artistic distortions or warping effects.

Shearing

- Practical

Affine Transformation

- Affine transformations modify the spatial arrangement of pixels in an image by:
  - Scaling: Uniformly or non-uniformly changing the size of objects.
  - Translation: Shifting objects to different positions.
  - Rotation: Rotating objects around a fixed point.
  - Shearing: Tilting objects along an axis, causing a slanted distortion.

Affine Transformation: Properties

- Line preservation:
  - Lines in the original image remain lines after the transformation, although their length and orientation might change.
- Parallel lines:
  - Parallel lines in the original image remain parallel after the transformation.
- Ratios preserved:
  - Ratios of distances between points are preserved, but not necessarily actual distances.

Affine Transformation

- Practical

Thank you

@mitu_skillologies

@mITuSkillologies

@mitu_group

@mitu-skillologies

@MITUSkillologies

kaggle

@mituskillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

@mituskillologies

contact@mitu.co.in

tushar@tusharkute.com