Chp-5

Date : / /
Page No. :

Computer Vision

→ Ability of computers to understand and analyze visual content in the same way as humans do.
→ This includes tasks such as recognizing objects, faces, reading text and understanding the context of an image or video.

→ Computer vision is closely related to AI.
→ ML algorithms are used to "train" a computer to recognize patterns and features in visual data, such as edges, shapes and colour.
→ Once trained, the computer can use this knowledge to identify and classify objects in new images and videos.

⇒ Working

→ Computer vision system consists of two main components

Sensory device	Interpreting device
- camera	- computer
- captures visual data from the env	- processes the data to extract meaning.

→ ~~Computer vision algorithms process images by looking for patterns in the shape, colors and textures of obj.~~

→ Computer vision algorithms ~~are based on~~ process images by looking for patterns in the pixels that make up the image.

→ These patterns can be used to identify and classify different objects in the image.

→ To analyze, algorithm first converts the image into set of numerical data that can be processed by the computer. Typically done by dividing the image into a grid of small units called pixels and representing each pixel with set of numerical values that describe its color & brightness.

→ once converted into numerical data, algorithm analyze it. Analyzed using techniques from ML to AI to recognize patterns in data and making decision based on those patterns.

Date :
Page No. :

⇒ Real-life Working of Computer Vision

1. IBM's use of computer vision to create "My Moments" for the 2018 Masters golf tournament.

2. Disney theme parks used this technology to improve operations.
If an attraction is experiencing technical issues, the system can predict the problem and automatically dispatch maintainence staff to fix it.

3. Google translate uses a smartphone camera & computer vision algorithms to analyze and translate text

⇒ Application in Real-world :

1. Facial recognition
- Used in security system
- Unlocking smartphones
- identifying individuals in photos.

2. Medical Imaging
- Assisting doctors in diagnosing diseases from x-rays, MRIs and CT scans.

Date :
Page No. :

3. Autonomous Vehicles.
- Enabling the cars to perceive their env,
- detect obstacles and making driving decisions

4. Surveillance
- monitoring and analyzing video feeds for security and crime prevention.

5. Agriculture
- Identifying crop disease
- estimating yields.

6. Retail
- Improving customer experiences with automated checkout and inventory management.

7. Augmented Reality:
- enhancing real world envs with computer-generated information, as seen in mobile apps and gaming.

⇒ Challenges of Computer Vision

1. Data limitations
   - → Computer vision requires a lot amount of data to train and test algorithm.
   - → This can be problematic in situations where data is limited or sensitive.

2. Learning rate
   - → Another challenge in computer vision is the time and resources required to train algo.
   - → While error rates have decreased over time, they still occur and it takes time for the computer to be trained to recognize and classify objects and patterns in images

3. Hardware requirements
   - → Computer Vision algorithms are computationally demanding, requiring fast processing & optimized memory architecture for quicker memory access.

4. Inherit complexity in the visual world.
   - → According to various conditions, there are an infinite number of possible scenes in a true vision system.

Date :
Page No. :

Geometric Primitives

→ Geometric Primitives are fundamental shapes and objects used in computer vision to represent and analyze visual data.
→ They serve as basic building blocks for more complex geometric structures.

→ includes :-

1. Points : points represent individual location in 2D or 3D space.
   - Used to mark key locations or images, such as corners of objects.

2. Lines : lines are used to represent straight or curved segments connecting two or more points.
   - They can be essential in tasks like line detection and boundary extraction

3. Curves : Curves represents more complex shapes ~~with~~ ~~stra~~ that are not limited to straight lines.
   - Examples :- circle, ellipse
   - Used in shape recognition

4 Polygons : Polygons are closed shapes with straight sides and defined vertices.
   - Use to represent & analyze complex objects.

Date : / /
Page No. :

→ Geometric primitives play a crucial role in various computer vision such as edge detection, shape recognition and object tracking.

2D and 3D Transformation

⇒ 2D Transformation

→ involve modifying the position, orientation and scale of objects or images in two-dimensional space.
→ These transformation are essential for tasks that deal with flat or planar object.
→ Common 2D Transformation include :-

1 Translation
- Translation moves an object / image in a particular dirⁿ by a specified distance.
- In 2D, it involves every point of the image by a fixed amount in the x and y dirⁿ.
- Used for moving objects on computer screen or aligning image.

2. Rotation
- Rotation changes orientation of the object / image about a fixed point, known as center of rotation.
- Used for image rotation or aligning objects in a scene.

Date :
Page No. :

3. Scaling
- Scaling alters the size of an object / image.
- It can make an object / image larger (scaling up) or smaller (scaling down).
- Scaling factors are used to control the amount of scaling in the x and Y dir^n.
- Used for resizing images, adjusting the proportions of objects, zooming in or out.

4. Shearing
- Shearing distorts the shape of an object / image by skewing it along one of the axes.
- Used for correcting perspective distortion in images or deforming objects.

⇒ 3D Transformation

→ Similar to 2D transformation but operate in a three-dimensional space, making them a applicable to 3D objects and scenes.
→ This transformation enables the manipulation and positioning of objects within a 3D env.
→ Common 3D transformations are :-

1. Translation
- 3D Translation involves moving an object in 3D space along with the x, y and z axes
- It shifts every point of object by specific distance along these axes

Date :
Page No. :

→ Translation is fundamental for relocating objects in 3D scenes.

2. Rotation
- In 3D, rotation can occur around any axes, not just a single point as in 2D.
- Euler angles are often used to describe 3D rotation
- Used to change orientation of objects, cameras, viewpoints in 3D space.

3. Scaling
- 3D scaling changes the size of an object along the x, y and z axes independently.
- It can make an object larger or smaller in one or more dimensions.
- Used for adjusting the proportion of objects in a 3D scene

4. Shearing
- 3D shearing involves skewing an object / parts of an object in 3D space,
→ Use to create distortions in 3D models / scenes.

⇒ 2D & 3D transformations are fundamental for computer vision involving in image manipulation, alignment and object positioning.

Date :
Page No. :

x. Orthographic Projection & Perspective Projection

⇒ This technique are used in computer vision and computer graphics to represent three dimensional (3D) objects and scenes in a two-dimensional (2D) space, such as on a computer screen or image.

⇒ Orthographic Projection

→ Here, 3D objects are projected onto a 2D plane without accounting for perspective.
→ This means that objects are represented in a way that preserves the relative sizes and angles of their parts, and parallel lines in the 3D world remain parallel in 2D projection.

→ Characteristics :-

1. Parallel Projections
 - lines that are parallel in 3D space remains parallel in 2D projection.
 - Thus provide accurate representation of shapes and relative sizes.
 - Useful in engeneering task & architectural design due to this

2. No perspective Distortion
- does not diminished sized of objects as they move farther from viewer

3. Use Cases
- technical drawings
- engineering design
- Computer aided design (CAD) applications

⇒ Perspective Projection

→ It simulates how objects appear in real world with depth perception.
→ It takes into account the converging lines and diminishing size of objects as they move farther from the viewer.

→ Characteristics :-

1. Converging lines
- 3D space appear to converge towards a single vanishing point in the 2D projections
- This creates a sense of depth & realism

2. Size Diminishing
- objects that are farther away from the viewer appear smaller in 2D projection, which accurately represents how we percieve objects in real world.

3. Realistic Depth
- It provides more realistic representation of 3D scenes, making it suitable for application where depth perception is crucial such as VR.

4. Use Cases
- Computer graphics for rendering 3D scenes
- 3D scene reconstruction
- Augmented reality