Task 10

- Registration/Referral id : SIRSS2218
- Full name : Shrikrushnakumar Bhausaheb Sondge

1. What is Pose Estimation?

Pose Estimation is a general problem in Computer Vision where we seek to detect the position and orientation of an object. This is typically done by identifying, locating, and tracking the number of key points on a given object or person. For objects, this could be corners or other significant features. And for humans, these key points represent major joints like an elbow or knee.

2. Pose Estimation Categories

When working with a person, the key points are defined (joints and landmarks) such as elbows, knees, neck, shoulder, hips, chest, etc. This can be categorized as **human pose estimation**. In the case of immovable or inanimate objects, the prediction is categorized as **rigid pose estimation**. This is because regardless of the orientation of the objects the distance between the key points does not change.

Based on what the input data is, the estimation can also be categorized as **2D**or **3D**pose estimation.

There also is a distinction between detecting one or multiple objects in an image or video. These can be referred to as **single** and **multi**-pose estimation

3. Pose Estimation Approaches

A simple way is to detect the person first, then estimate the key points, and lastly estimate the pose for the person. This constitutes a top-down approach.

Another way is to detect the key points first, then group the key points and associate them to a distinct person. This method is known as the bottom-up approach.

4. Challenges

There is a huge variety of challenging difficulties concerning pose estimation.

First of which is a clear image, more often than not the images available are not clear, making it difficult to identify the subjects.

Secondly, all the parts of the body that we use for the key points should be visible. In cases of low-resolution images, this task becomes difficult. Also, in the case of multiple people present in the image, this becomes an even more difficult task.

In the case of Multi pose estimation, there is the problem of overlapping. Usually, we can see multiple people, and the key points may overlap.

With the advent of smartphones and their cameras, there is another problem with the images being of varied sizes. The rescaling or resizing is another challenge where we ought to maintain the details while making sure that the model can precisely extract the relevant features.

5. Use cases and applications

A. Human activity and movement:

Tracking the variations in the pose of a person over a period can also be used for activity, gesture, and gait recognition. Some applications could be as below:

- Application to identify if a person follows the exercise regime properly

- Application to identify the health of a person – In case the person collapses inform somebody close or paramedics

- Application to identify body language and flag suspicious individuals to nearby authorities - used at airports

B. Augmented reality

- CGI Application to track the human pose variations to render graphical animations – e.g: Thanos

C. Animation & gaming

- Identify and track movements in Gaming

D. Robotics

- Robots can be taught to mimic human poses, activities by tracking and following human instructor demonstrations, instead of manually programming robots