

ASSIGNMENT – 10

Name :- Sanjay Dey

Ref ID :- SIRSS2204

POSE ESTIMATION USING COMPUTER VISION

Pose estimation is a computer vision technique that is used to predict the configuration of the body(POSE) from an image. The reason for its importance is the abundance of applications that can benefit from technology.

Human pose estimation localizes body key points to accurately recognize the postures of individuals given an image. These estimations are performed in either 3D or 2D.

The main process of human pose estimation includes two basic steps:

- i) localizing human body joints/key points**
- ii) grouping those joints into valid human pose configuration**

What is OpenCV ?

OpenCV Python is a library of Python bindings designed to solve computer vision problems. It mainly focuses on image processing, video capture and analysis including features like face detection and object detection.

Use of OpenCV in python :

OpenCV Python is nothing but a wrapper class for the original C++ library to be used with Python. Using this, all the OpenCV array structures gets converted to/from NumPy arrays. This makes it easier to integrate it with other libraries which use NumPy. For example, libraries such as SciPy and Matplotlib.

What is human pose estimation?

Human pose estimation aims at predicting the poses of human body parts and joints in images or videos. Since pose motions are often driven by some specific human actions, knowing the body pose of a human is critical for action recognition.

What is 2D Human Pose Estimation?

2D human pose estimation is used to estimate the 2D position or spatial location of human body key points from visuals such as images and videos. Traditional 2D human pose estimation methods use different hand-crafted feature extraction techniques for the individual body parts.

Early computer vision works described the human body as a stick figure to obtain global pose structures. However, modern deep learning based approaches have achieved major breakthroughs by improving the performance significantly for both single-person and multi-person pose estimation. Some popular 2D human pose estimation methods include OpenPose, CPN, AlphaPose, and HRNet (we will cover them and others later in this article).

What is 3D Human Pose Estimation?

3D Human Pose Estimation is used to predict the locations of body joints in 3D space. Besides the 3D pose, some methods also recover 3D human mesh from images or videos. This field has attracted much interest in recent years since it is used to provide extensive 3D structure information related to the human body. It can be applied to various applications, such as 3D animation industries, virtual or augmented reality, and 3D action prediction. 3D human pose estimation can be performed on monocular images or videos (normal camera feeds).

Human body modeling

In human pose estimation, the location of human body parts is used to build a human body representation (such as a body skeleton pose) from visual input data. Therefore, human body modeling is an important aspect of human pose estimation. It is used to represent features and keypoints extracted from visual input data. Typically, a model-based approach is used to describe and infer human body poses and render 2D or 3D poses.

There are three types of models for human body modeling:

- Kinematic Model
- Planar Model
- Volumetric model

How does Pose Estimation work?

Pose estimation utilizes pose and orientation to predict and track the location of a person or object. Accordingly, pose estimation allows programs to estimate spatial positions (“poses”) of a body in an image or video. In general, most pose estimators are 2 steps frameworks that detect human bounding boxes and then estimate the pose within each box.

Pose estimation operates by finding key points of a person or object. Taking a person, for example, the key points would be joints like the elbow, knees, wrists, etc. There are two types of pose estimation: multi-pose and single pose. Single pose estimation is used to estimate the poses of a single object in a given scene, while multi-pose estimation is used when detecting poses for multiple objects.

Human pose estimation on the popular COCO Dataset can detect 17 different keypoints (classes). Each key point is annotated with three numbers (x, y, v), where x and y mark the coordinates, and v indicates if the key point is visible.

Pose Estimation with Deep Learning

With the rapid development of deep learning solutions in recent years, deep learning has been shown to outperform classical computer vision methods in various tasks, including image segmentation or object detection. Therefore, deep learning techniques brought significant advances and performance gains in pose estimation tasks.

Next, we will list and review the popular pose estimation methods.

The Most popular Pose Estimation methods

- Method #1: High-Resolution Net (HRNet)
- Method #2: OpenPose
- Method #3: DeepCut

- Method #4: Regional Multi-Person Pose Estimation (AlphaPose)
- Method #5: Deep Pose
- Method #6: PoseNet
- Method #7: Dense Pose

Use Cases and Applications of Pose Estimation

Human pose estimation has been utilized in a wide range of applications, including human-computer interaction, motion analysis, augmented reality, and robotics.

Pose estimation has applications in lots of fields, some of which are listed below:

- 1) Training Robots – Robots can be taught to mimic human poses, activities by tracking and following human instructor instead of manually programming robots.
- 2) Animation and gaming – Identify and track movements in gaming.
- 3) Human activity and movement – Tracking the variation in pose of a person over a period of time.
- 4) Augmented reality – GCI applications to track the human pose variations to render graphical animations.
- 5) The Covid-19 quarantine has resulted in the surge of the use of home exercise equipment. One such product is the smart mirror, such as the ones made by [MIRROR](#), which streams exercise classes directly to your mirror!