Human Pose Estimation

Flavio Amurrio-Moya, G00593001

Octobe 2021

1 Introduction

In recent years, the application of Human Pose Estimation has progressed exponentially in terms of speed and accuracy. From being able to detected simple 2D poses of one person at the time to real time multi person detection in videos, this rapid evolution is mainly thanks to the advances of Deep Neural Networks, which help with the detections on body part sections and joints within the human body. In lieu of this, [...](topic) began originally as a means to [...] before its broader application into healthcare, entertainment, and videography in general. This document will seek to explain Human Pose Estimation(topic)'s history, its conception and progression according to the latest research, as well as its shortcomings and controversy stirred by its impact on the discipline of [...] and [...]. The origin of Pose Estimation was first brought about by [...] in [...]. During this time, there was a strong need for [...] which prompted hastened research in [...] which was an already emerging topic because of [...]. With [...](originator)'s idea/With the idea of [...], [...](topic) would be able to resolve the shortcomings of [...], and thus [...](topic) began to gain interest and influence. Perhaps the greatest influence (and possible funding [...](citation)) came from the entertainment industry. As more and more realistic forms of cinematography advanced in the late 1990s and early 2000s [...](citation), so to would these require more intricate and complex renditions of the human anatomy. This would later evolve beyond the scope of mere appearance to human behaviour and movement as well as we delve into the more technical aspect in the below survey.

2 Survey

To achieve a realistic depiction of the human form, Human Pose Estimation requires, in broad terms, two main steps - preprocessing and body parts parsing. We will explore each step and its subsequent processes in detail, beginning with preprocessing.

2.1 Prepocessing

Preprocessing can be generously defined as [...]. In this discipline in particular, it involves the initial division of the human body into several parts (known as segmentation)[...](citation), of taken with the exact measurements as expressed by a camera's calibration. The camera, in this instance, requires varying levels of perspective for it's beginning depiction, before it can be compiled through a process known as [...]. It does not appear that a specific brand of particular camera or video recorder is needed, nor does there seem to be a specific amount of differing perspectives needed before the dataset can be considered sufficient enough for processing [...](citation). A popular body capturing device found through research gathered appears to be the usage of Kinect [...](citation) - most likely due to its dominating presence of the market and cost [...](citation). After the adequate amount of data is gathered, it is later compiled into a [...] by [...]. The purpose of this compilation is so that [...] can determine [...]. This is typically an automated process performed by [...] which [...] by [...]. After the fundamental skeleton has been rendered, another proponent of preprocessing involves body localization []. Body localization, as the name suggests, involves using a [...] to recognize and locate a humanoid figure in what can be considered a heavily convoluted image. Perhaps a robust example of this could be asking an [...] to distinguish a human body while it stands in front of a painted mosaic. The benefit of this technology has evolved to be used as [...], and under the context of HPE (Human Pose Estimation) it also helps to define the complete range and [...] of the body as it is expressed on said media. This flows into the usage of Human detection, which involves a similar concept in the recognition of a person's location and space in a given image. Human detection is done by [...] and later involves [...], similar to [...]. Note the primary difference between human detection and body localization as this: while [...] is [...] and can detect [...] based on [...], [...] captures [...] by evaluating [...]. While their method of capture may be a bit different, the two processes

do share the result that comes about after they've performed their relevant calculations. However, the main goal of [...] is the [...] of [...]. Perhaps the greatest use is for [...], of which [...] could possibly for short on, thus making [...] the popular choice.

The final method of preprocessing involves Background Subtraction. Background Subtraction involves [...]. It appears to be different from [...] proves unable to [...] due to [...]. The process requires [...] to later [...] into a [...] which provides an image with [...]. This process can be improved upon by [...]. Background Subtraction is particularly useful for [...]. However, a main weakpoint is that it [...], this is due to [...].

Note that the above processes are components of preprocessing, but are not inherent steps required in any particular order (at least according to the resources gathered). It is an educated estimation that while each method contains its strengths and weaknesses, a combination of the methods summarized above would lead to the greatest dataset needed to impose [...]. When attempting to replicate [...], it is important to note that each process required the same need for [...], this thus acts as a cornerstone for preprocessing as well as [...] as a whole, as it contributes to [...] in detail.

I would like to explore the field Human Project proposal on Human Pose Estimation Define what it is; in layment terms is the process of identifyind the pose of a person based on still images or video. These include 2d and 3d poses. IN the video/game media, people have use motion caspturing systems that rely on th subject to markes on their body to better track the individuals movements. This can be expensive since a whole setup sis required from special suuits to unobstructed backgrounds Other solutins include the use of infrared dots being project onto the indfivudal and then using a camara to capture these points, this was most publicaly used by microsfot Kinect.

Body part/ join detection Different areas 2d single person, multiperson in imgaes and Videos

Applications: Robotics Motion capturing for digital entertainment Virtual./Augment Reality Medical Excersise monitoring Body language analysis Security/Instruction detection Humcan computer interaction Medical monitoring

process Body/human/object detection: Being able to detect where the substect is and then being Remove noise: Background substraction. Isolating the subject is key in order to properly analyze and determine the pose

2 - 3 sentences: How does it work?

2 - 3 sentences: How does it work?

2 - 3 sentences: How were these achieved Results Weakness of study

- Body parising strongpoints - Weakness of Ai Evolution and predopistion to dataset used - Controversy and hinderance of study

Scoring: Problems: background context obstruction

limitations

Future

3   Approach

Write about the proposed work touching the research gap you are addressing, novelty etc. When you are using some references, make sure you list them in the reference below and cite them in this section [3] [2].

And in the last paragraph write about the proposed deliverable aimed from this thesis [4].

Proposed Guides: Dr./Mr./Ms Name of faculty, Professor/Associate/Assistant Professor, Department, Campus.

- Well organized

- Well written

- Ideas are clearly stated and the task of the project is well-defined

- Concepts are formally stated

- Statements are correct

- Broad and deep coverage of the literature

- Critical analysis of the literature

- Pros and cons, gaps, limitations, drawbacks of existing work are identified

- Connections and relationships among the various methods are identified and discussed

- The writing shows that the student has a good understanding of existing work

- Use of tables and graphics to highlight the features of the methods being discussed helps the reader to grasp the overall picture and is highly recommended

- Latex is required for writing the proposal

References

[1] N.J. Higham. *Handbook of writing for the mathematical sciences.* Society for Industrial Mathematics, 1998.

[2] J. M. Linhart. Algorithm 885: Computing the logarithm of the normal distribution. *ACM Transactions on Mathematical Software*, 35:Article 20, 2008.

[3] Jean Marie Linhart. Teaching writing and communication in a mathematical modeling course. *PRIMUS*, 24(7):594–607, 2014.

[4] Wikipedia. Fractal, 2012.