Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation

Maksym Andriushchenko

2565540
&
FAN Yue
2564216

Saarland Univ.

Abstract

We address the problem of single person pose estimation from 2D images. Our method is aspired from the paper Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation, the model is a combination of a CNN and a PGM trained end-to-end jointly. We reproduce the same results of the authors and explore that with Batch Normalization, we achieve a faster convergence.

1. Problem Overview

2. Methods

3. Evaluation

4. Future work

References

[1] Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation

[2] Learning Human Pose Estimation Features with Convolutional Networks

[3] Multimodal decomposable models for human pose estimation

Question: *Find the optimal alignment between two ellipses. Debug until you achieve a good result!*

It is already optimal. We only added scaling by multiplying point cloud by a scalar, which lead to the same scaling across all coordinates.

Question: *Try different ellipse rotations and scales. Does the method find the optimal solution for any scale, rotation and translation?*

Yes. We can consider a few extreme cases and judge according to the residuals:

- Scaled ellipse: scale=4, rotation=0, translation: (0, 0) => residuals 0.
- Translated ellipse: scale=1, rotation=0, translation: (2, 4) => residuals 0.
- Rotated ellipse: scale=1, rotation=pi/4, translation: (0, 0) => residuals 0.
- All transformations: scale=3, rotation=pi/4, translation: (2, 4) => residuals 0.

Since residuals are always 0, we can conclude that we have a perfect fit in each case.