Vincent Matthys
ENS Paris Saclay

Automatic estimation of 3D pose from 2D ground truth poses using videos

Vincent Matthys
vincent.matthys@ens-paris-saclay.fr

I decided to work on my one on project D named : **Automatic estimation of 3D pose from a single image**. The plan of the work is the following :

- Reproduce results obtained in [1] using only 20 epochs with a GTX-1050 4Go, with ground-truth detection, focusing only on the 2D ground-truth to 3D positions problem.
- Using the simple model proposed in [2] for modeling human action using video information through sequence-to-sequence, implement a sequence-to-sequence model to use video information in Human3.6M database.
- Compare the previous resutls
- If possible, use combine both model as suggested in [3]
- Test on the shorter manipulations videos dataset, and characterize the working conditions and failures modes

Références

[1] J. Martinez, R. Hossain, J. Romero et J. J. Little, "A simple yet effective baseline for 3d human pose estimation", *arXiv preprint arXiv :1705.03098*, 2017.

[2] J. Martinez, M. J. Black et J. Romero, "On human motion prediction using recurrent neural networks", *arXiv preprint arXiv :1705.02445*, 2017.

[3] M. R. I. Hossain et J. J. Little, "Exploiting temporal information for 3D pose estimation", *arXiv preprint arXiv :1711.08585*, 2017.