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

Y. Fan & M. Andriushchenko

Saarland Univ.

30 January 2018

What is state-of-the-art in **human pose estimation**?

We all know that CNNs show state-of-the-art in many computer vision tasks.

But what about PGMs? Do we actually need them?

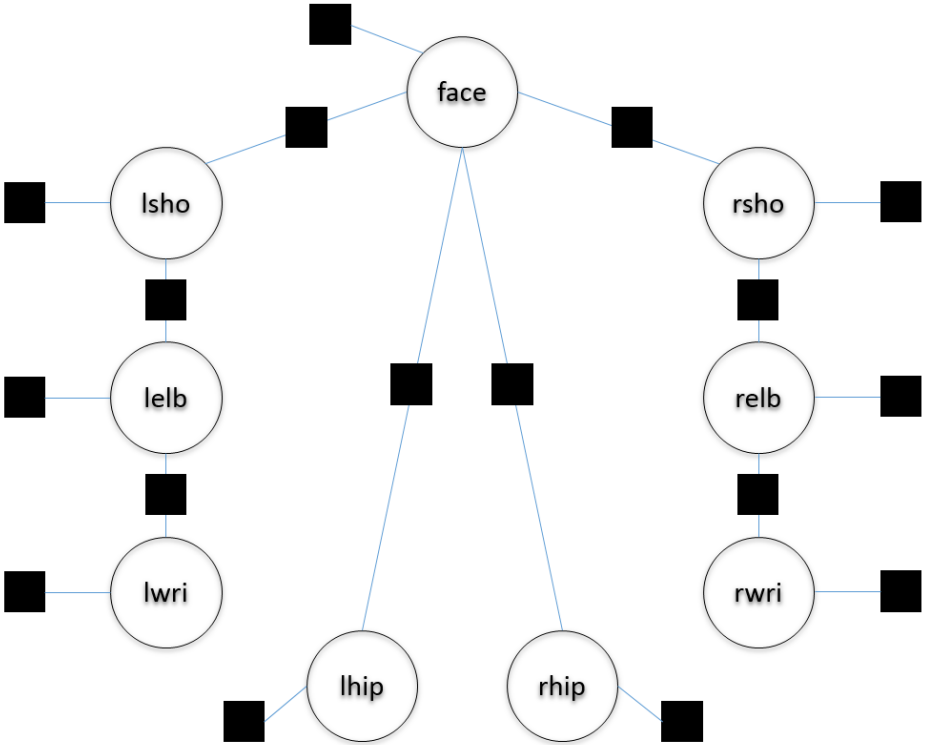
Higher-Level Spatial Model

Problem: Part Detector produces many false-positives.

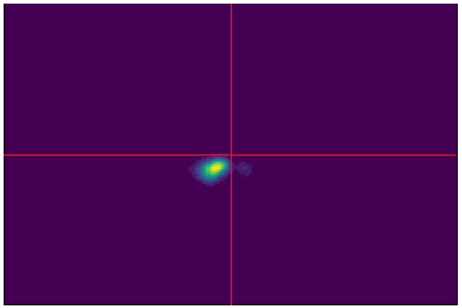
Solution: use a Spatial Model to enforce the consistency.

Spatial Model as a PGM

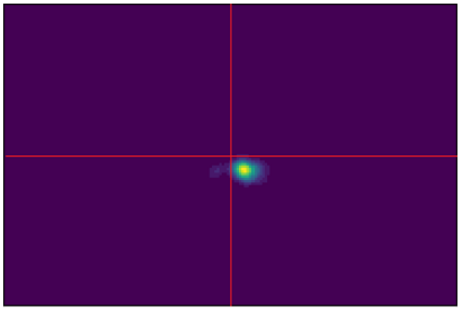
We adopt the star model.



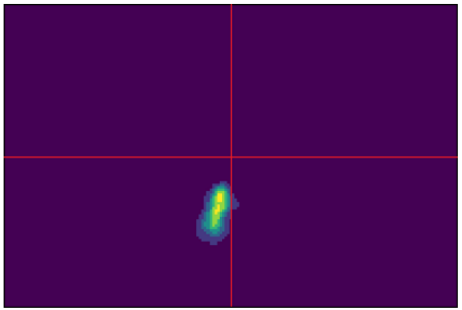
Pairwise Potentials



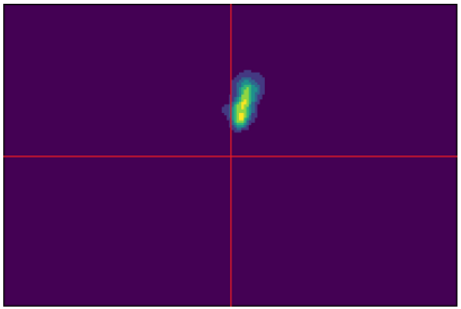
Left shoulder given face



Right shoulder given face

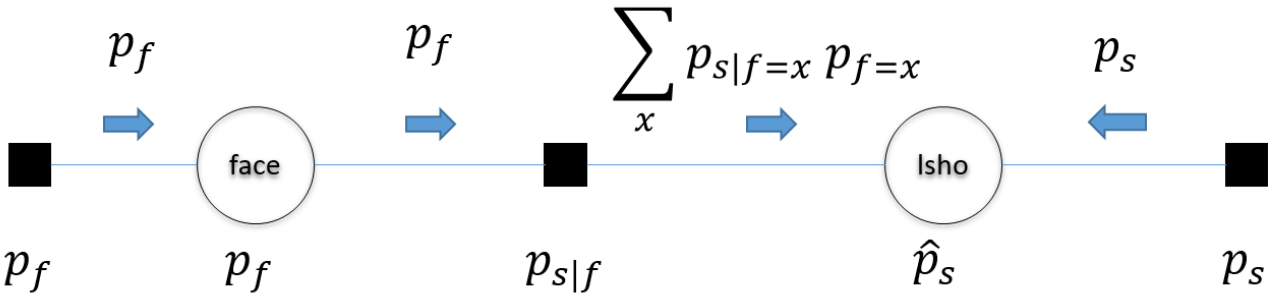


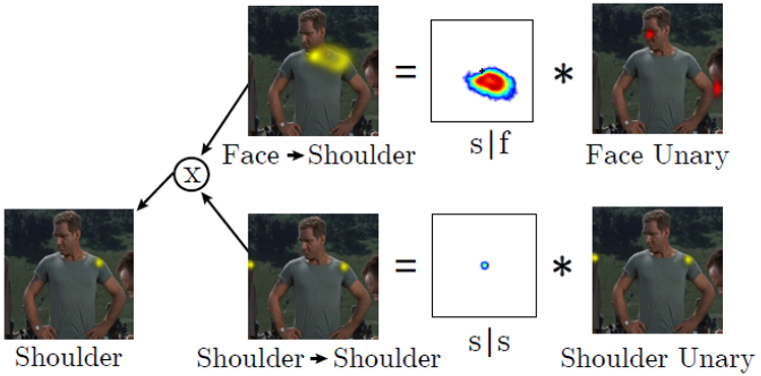
Left hip given face



Face given left hip

Inference on PGM





$$\hat{p}_i \propto p_i \prod_{u \in U} (p_{i|u} * p_u)$$

where U is a set of neighbouring nodes of body part i

Spatial Model as a trainable PGM

The Spatial Model can also be modeled as fully connected graph with trainable parameters.

Star PGM:

- computationally efficient (during the train phase).
- Less parameters to be train.
- Inference is exact.

Fully Connected PGM:

- More model capacity.
- **The model is learned from the data, no need of expert prior.**
- Loopy structure has no guarantee of convergence.

Conclusions

- We open sourced all our code in our Github repository: `https://github.com/max-andr/cnn_mrf_hybrid_for_hpe`!
- Up to our knowledge, this is the first implementation of the presented paper [1].

Thanks for your attention!

Any questions?

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

[2] Learning Human Pose Estimation Features with Convolutional Networks