Face Rectification on Feature Maps for Recognition

Peng Lu

SIGS, Tsinghua University, China

Abstract

- **Face rectification with image warping destroys local texture while increasing the semantic alignment.**
- **Each pixel on feature map describes a patch on original image. It saves texture information in channel dimension.**
- **Warping on feature maps is less harmful to texture information, which is important for recognition.**

Motivation

Face rectification

- reduces geometric variation of face images → makes it easier for CNN to recognize profile faces.
- is realized via image warping, which relocate pixels according to their semantic. (Side effect: Image warping is destructive to texture, which is important to recognition.)

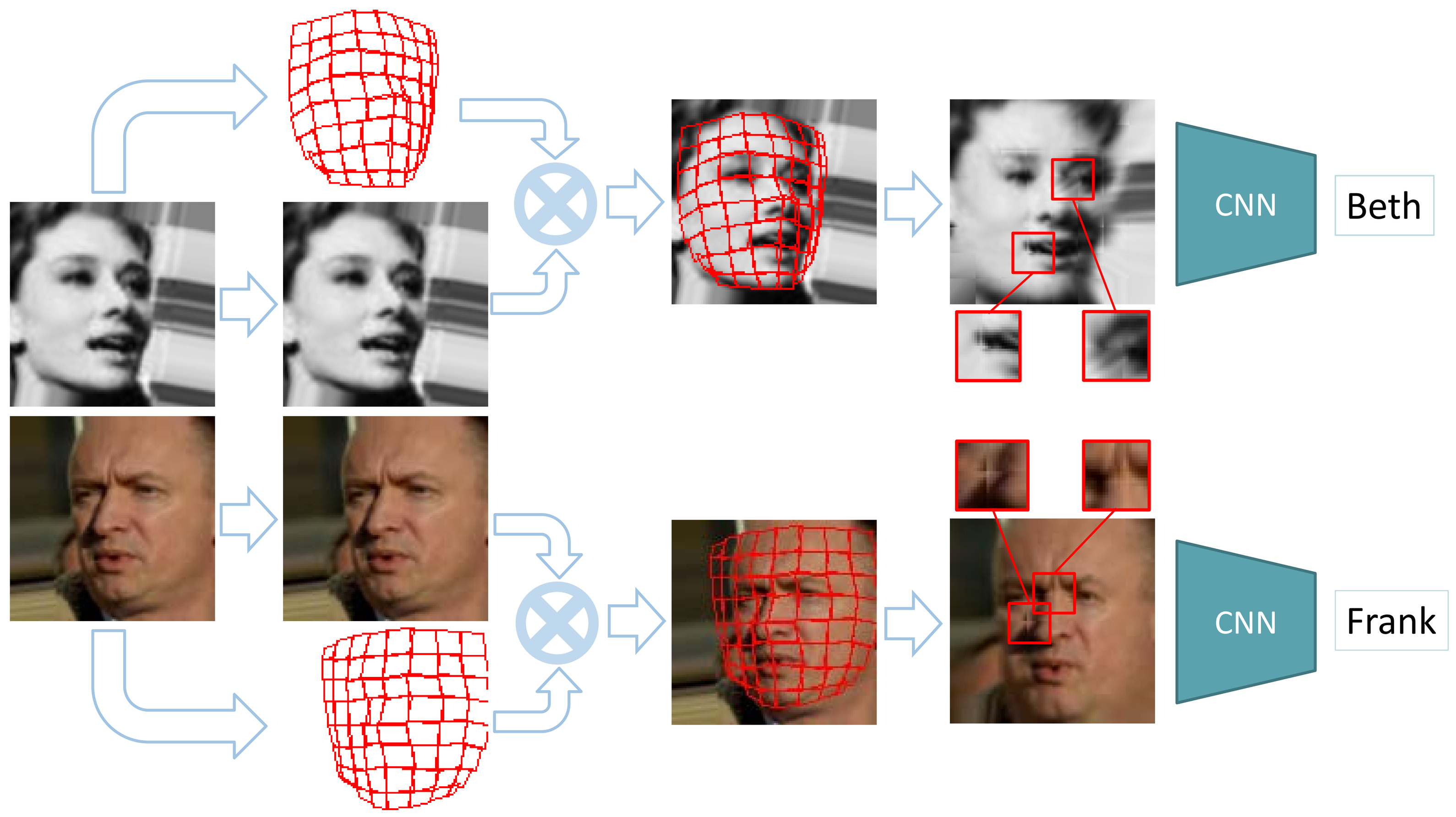

Figure 1. Face rectification on original images

On feature map, a pixel vector is computed from a patch on the original image and saves local texture information.

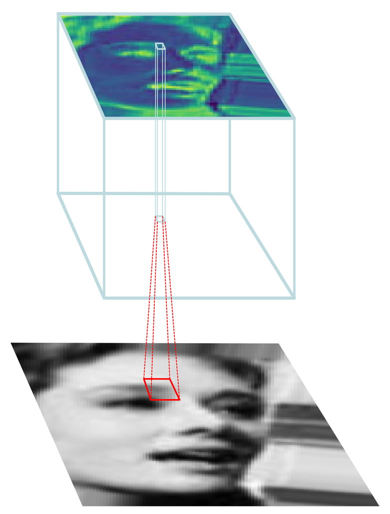

Figure 2. Semantic correspondence of feature map and original image

Method

- Conduct face rectification on feature maps.
  1. Compute transformation parameters from original images.
  2. Apply image warping on feature maps at a selected stage with the pre-computed parameters.
- Combined with various face rectification methods: GridFace[1], STN-TPS[2].

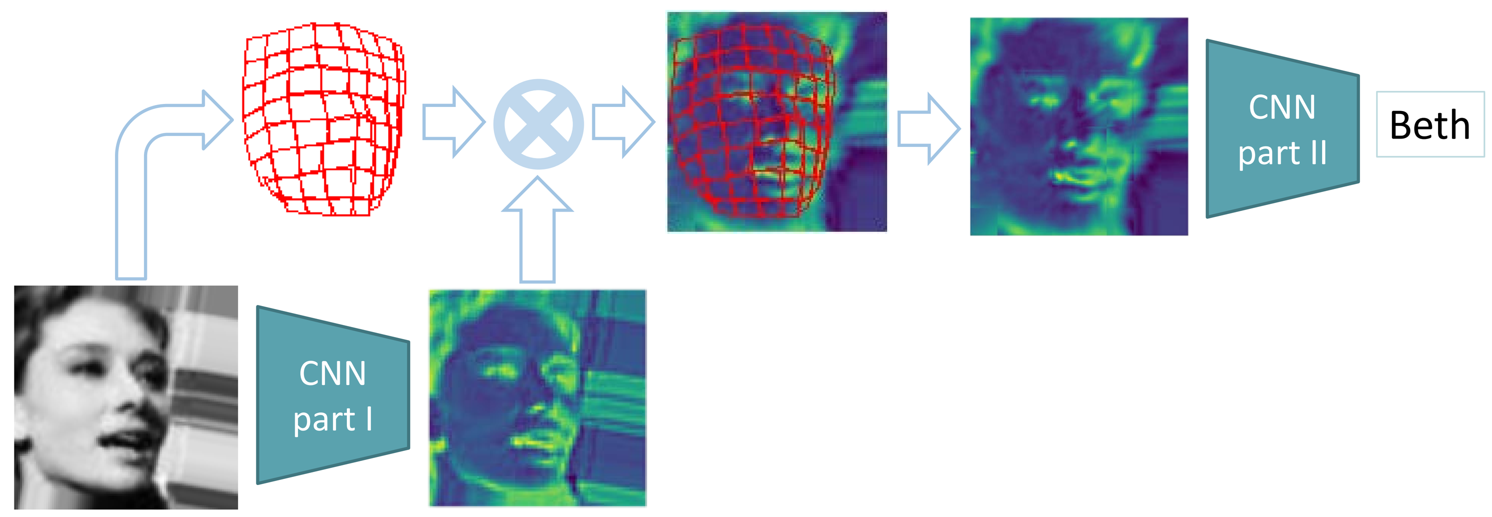

Figure 3. Face rectification on feature maps

Results

Method	MultiPIE			IJB-A
	60°	75°	90°	top-1
w/o face rect.	97.17	92.32	79.81	91.25
Grid@Img	97.99	94.86	86.77	91.45
Grid@Fmap-0	97.98	94.88	87.00	91.78
Grid@Fmap-1	**98.04**	**94.91**	**87.60**	**92.01**
TPS@Img	98.30	95.21	87.52	91.62
TPS@Fmap-0	98.39	95.26	87.48	92.09
TPS@Fmap-1	**98.62**	**96.00**	**88.92**	**92.58**
TPS@Fmap-2	98.00	94.71	86.65	91.40
TPS@Fmap-3	98.06	94.29	84.49	91.35

Table 1. Face identification accuracy.

- With both GridFace and STN-TPS, applying face rectification on feature maps at stage 1 is best in face recognition task.
- If convolution kernels in CNN part I are replaced with 1x1 conv, where pixels on feature maps no longer contain texture information, applying rectification on feature maps brings no performance boost.
- Conducting face rectification on feature maps also has side effects. I have not figured out why.

Method	MultiPIE			IJB-A
	60°	75°	90°	top-1
w/o face rect.	94.69	87.96	72.05	90.47
TPS@Img	**97.34**	93.30	**82.83**	**90.90**
TPS@Fmap-0	97.24	**93.43**	82.10	90.84
TPS@Fmap-1	96.97	92.96	80.98	90.57

Table 2. Performance with CNN that uses 1x1 convolutional kernels in input layer and 1^{st} stage.

Conclusion

- Conducting face rectification on feature maps at an appropriate stage can achieve better performance than on original images.
- Feature maps save texture information in channel dimension and thus are robust to spatial transformation (warping).

References

[1] Erjin Zhou, Zhimin Cao, and Jian Sun. Gridface: Face rectification via learning local homography transformations. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–19, 2018.

[2] Max Jaderberg, Karen Simonyan, Andrew Zisserman, et al. Spatial transformer networks. In *Advances in neural information processing systems*, pages 2017–2025, 2015.