# Pose-Guided Photorealistic Face Rotation

Yibo Hu[1,2], Xiang Wu[1], Bing Yu[3], Ran He[1,2], Zhenan Sun[1,2]

[1]CRIPAC & NLPR & CEBSIT CASIA, [2]University of Chinese Academy of Sciences, [3]Noah's Ark Laboratory, Huawei Technologies Co., Ltd.

CVPR 2018 SALT LAKE CITY • JUNE 18-22

## Background:

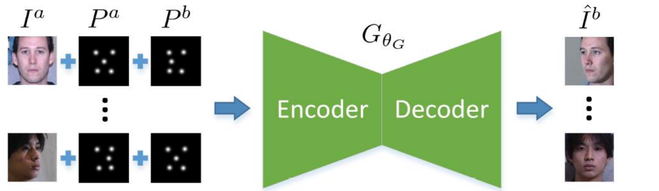
Face rotation provides a cheap but effective way for data augmentation and representation learning of face recognition. It aims to rotate a normalized face to arbitrary poses, where only yaw is considered.

## Contributions:

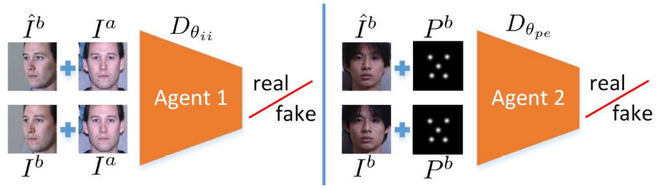
- A Couple-Agent Pose-Guided Generative Adversarial Network is proposed for face rotation from a single image in 2D space.
- Landmark heatmaps are used as controllable signals in pose-guided generator.
- The couple-agent discriminator combines prior domain knowledge of pose and local structure of face to reinforce the realism of synthetic arbitrary view faces.

## Framework:

Pose-Guided Generator




Couple-Agent Discriminator




## Loss Functions:

$$\min_{\theta_G} \max_{\theta_{ii},\theta_{pe}} L = \lambda_1 L_{pix} + \lambda_2 L_{adv}^{ii} + \lambda_3 L_{adv}^{pe} + \lambda_4 L_{ip} + \lambda_5 L_{tv}$$

$$L_{pix} = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{W_s H_s C}\sum_{w,h,c=1}^{W_s,H_s,C}\left|\hat{I}^b_{s,w,h,c} - I^b_{s,w,h,c}\right|$$

$$L_{adv}^{ii} = E_{I^b\sim P(I^b)}\left[\log D_{\theta_{ii}}\left(I^b, I^a\right)\right] + E_{\hat{I}^b\sim P(\hat{I}^b)}\left[\log\left(1 - D_{\theta_{ii}}\left(\hat{I}^b, I^a\right)\right)\right]$$

$$L_{adv}^{pe} = E_{I^b\sim P(I^b)}\left[\log D_{\theta_{pe}}\left(I^b, P^b\right)\right] + E_{\hat{I}^b\sim P(\hat{I}^b)}\left[\log\left(1 - D_{\theta_{pe}}\left(\hat{I}^b, P^b\right)\right)\right]$$

$$L_{ip} = \left\|D_{ip}^{p}(\hat{I}^b) - D_{ip}^{p}(I^b)\right\|_F^2 + \left\|D_{ip}^{fc}(\hat{I}^b) - D_{ip}^{fc}(I^b)\right\|_2^2$$

$$L_{tv} = \sum_{c=1}^{C}\sum_{w,h=1}^{W,H}\left|\hat{I}^b_{w+1,h,c} - \hat{I}^b_{w,h,c}\right| + \left|\hat{I}^b_{w,h+1,c} - \hat{I}^b_{w,h,c}\right|$$

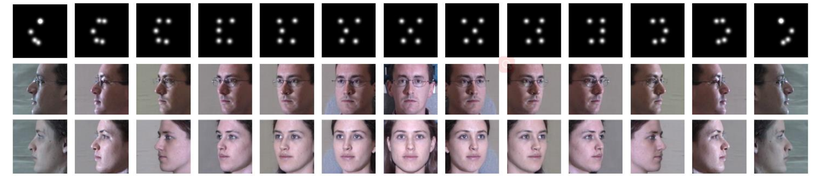
## Quantitive Results:

| Method | ±90° | ±75° | ±60° | ±45° | ±30° | ±15° |
|---|---|---|---|---|---|---|
| FIP+LDA[40] | - | - | 45.9 | 64.1 | 80.7 | 90.7 |
| MVP+LDA[41] | - | - | 60.1 | 72.9 | 83.7 | 92.8 |
| CPF[34] | - | - | 61.9 | 79.9 | 88.5 | 95.0 |
| DR-GAN[28] | - | - | 83.2 | 86.2 | 90.1 | 94.0 |
| FF-GAN[35] | 61.2 | 77.2 | 85.2 | 89.7 | 92.5 | 94.6 |
| TP-GAN[14] | 64.64 | 77.43 | 87.72 | 95.38 | 98.06 | 98.68 |
| Light CNN[29] | 5.51 | 24.18 | 62.09 | 92.13 | 97.38 | 98.59 |
| CAPG-GAN | **66.05** | **83.05** | **90.63** | **97.33** | **99.56** | **99.82** |

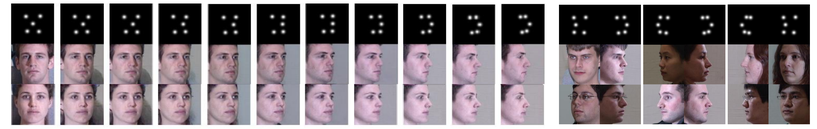
## Ablation Study:

| Method | ±90° | ±75° | ±60° | ±45° | ±30° | ±15° |
|---|---|---|---|---|---|---|
| w/o $L_{ip}$ | 40.77 | 46.13 | 53.25 | 64.74 | 76.16 | 86.12 |
| w/o $L_{tv}$ | 61.33 | 78.98 | 87.68 | 95.58 | 99.03 | 99.74 |
| w/o $L_{adv}$ | 46.83 | 56.90 | 67.68 | 85.68 | 96.26 | 99.50 |
| w/o $L_{adv}^{ii}$ | 54.68 | 66.09 | 75.90 | 89.38 | 97.79 | 99.73 |
| w/o $L_{adv}^{ip}$ | 57.78 | 71.17 | 82.05 | 92.50 | 97.57 | 99.63 |
| CAPG-GAN | **66.05** | **83.05** | **90.63** | **97.33** | **99.56** | **99.82** |


## Qualitative Results:



(a) +90° (b) +75° (c) +60° (d) +45° (e) +30° (f) +15° (g) 0° (h) −15° (i) −30° (j) −45° (k) −60° (l) −75° (m) −90°



(a) Source (b) 7.5° (c) 15° (d) 22.5° (e) 30° (f) 37.5° (g) 45° (h) 52.5° (i) 60° (j) 67.5° (k) 75° +30° → −60° +60° → −75° +75° → +30°