Pose-Guided Photorealistic Face Rotation

Yibo Hu

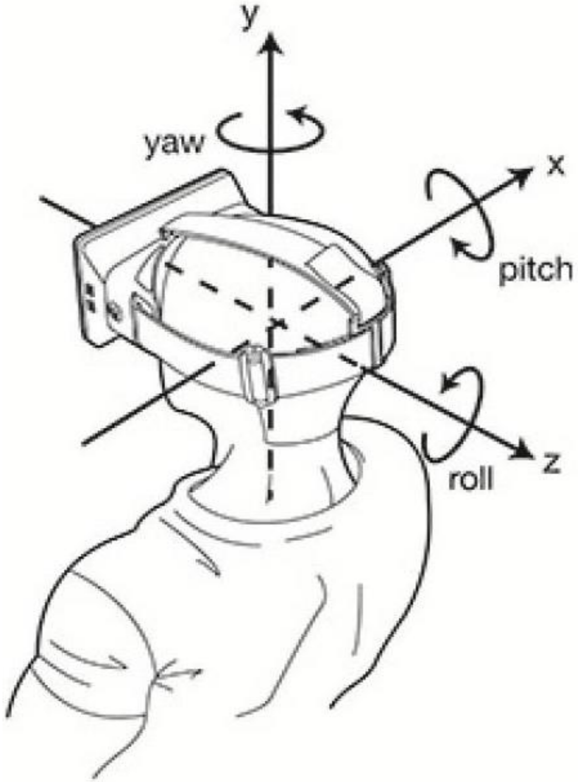
CRIPAC, NLPR

CASIA

2018-05-11

Face Rotation

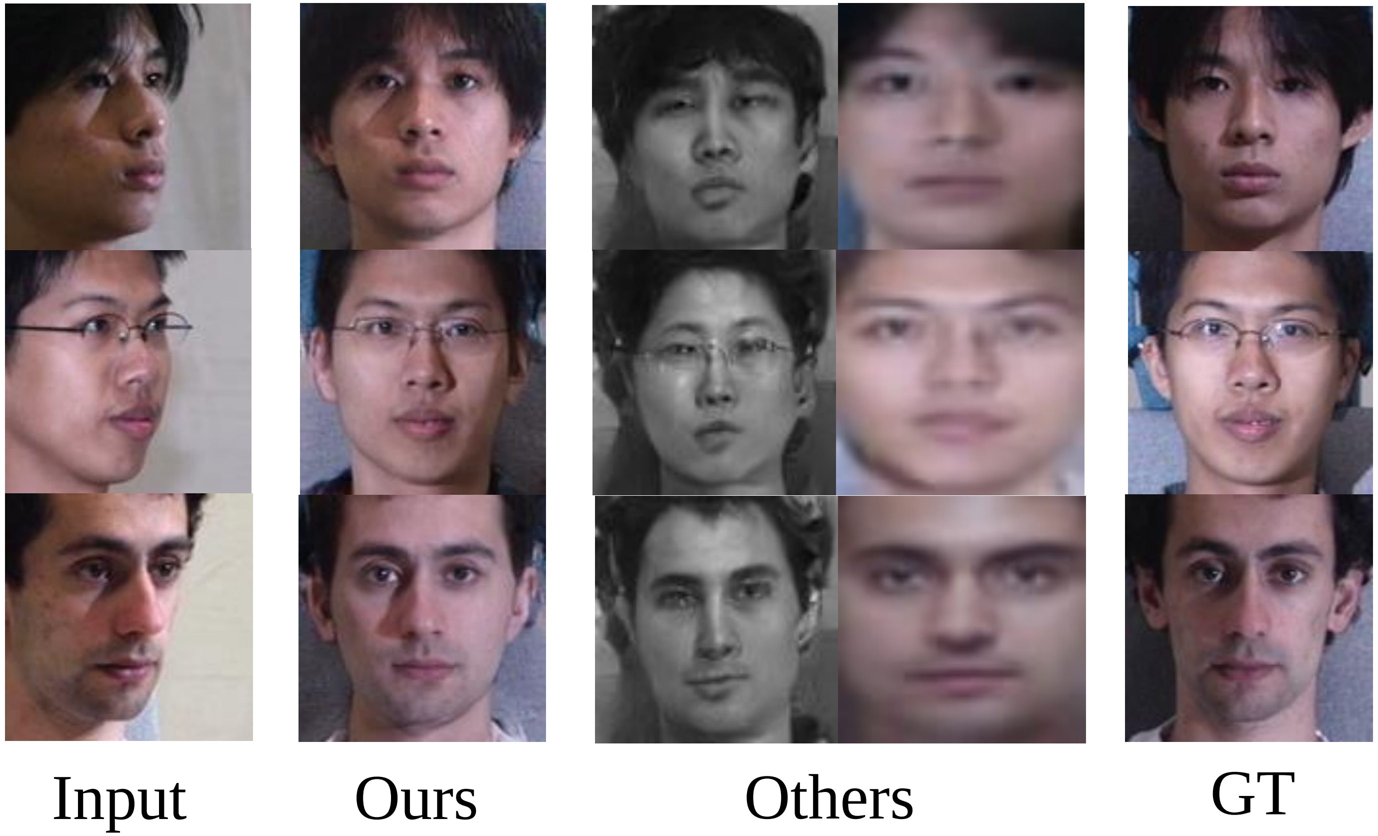
- Rotate a normalized face to arbitrary poses.
- Academia and industry pay close attention to it.
- Only yaw is considered.



Applications:

- Face edition
- Pose invariance
- Data augmentation
- Representation learning
- ……

Face Rotation



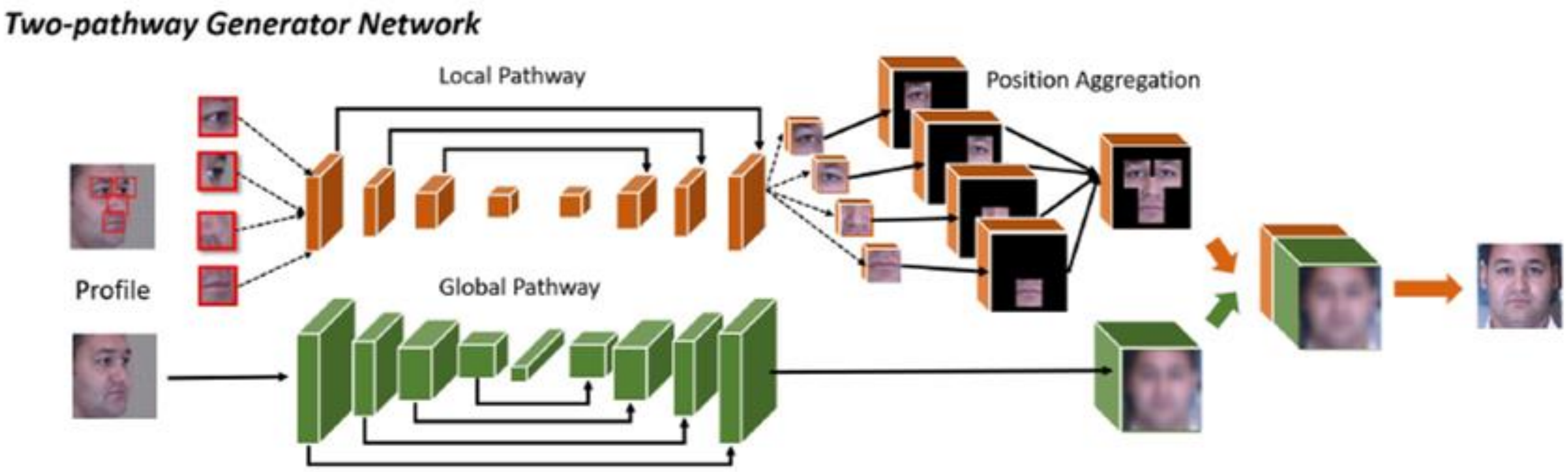
- Photo-realistic
- High-resolution
- Identity preserving
- Ill-posed problem

Related Work

- **Frontalization**
  - 2D methods
    - TP-GAN ICCV17
    - PIM CVPR18
  - 3D methods
    - Hassner et al. CVPR15
    - HPEN CVPR15
    - FF-GAN ICCV17

- **Rotation**
  - 2D methods
    - CPF $\pm 60^o$ CVPR15
    - DR-GAN $\pm 60^o$ CVPR17
    - $\boxed{\text{CAPG-GAN } \pm 90^o \text{ CVPR18 (Ours)}}$
  - 3D methods
    - DA-GAN $\pm 90^o$ NIPS17
    - UV-GAN $\pm 90^o$ CVPR18

Motivation — TP-GAN



Input

TP-GAN

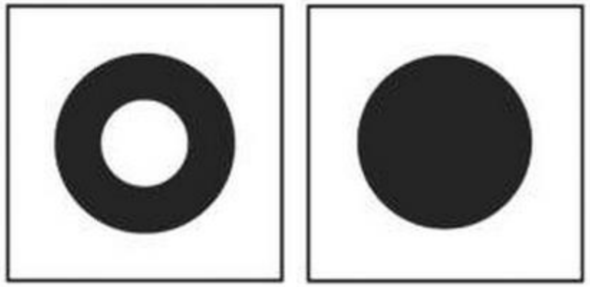
DR-GAN

- One global net and four local nets cause inference bottleneck.
- The architecture and loss designments are specific for frontalization.
- More flexible controls are needed for arbitrary pose synthesis.

TP-GAN: R. Huang, etal. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. In ICCV, 2017.

Motivation — Global-First Topological Perception

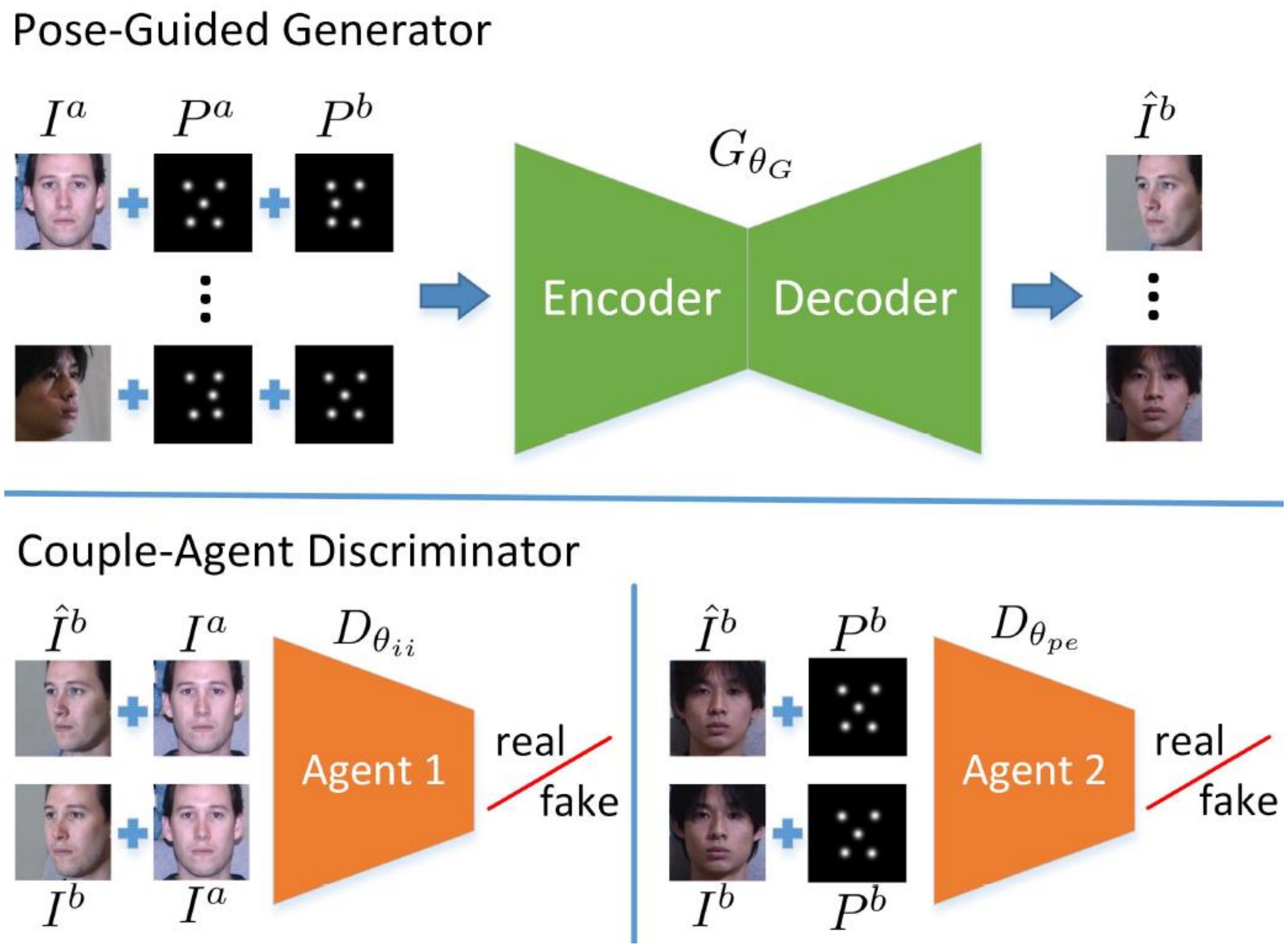
- “The visual system is sensitive to global topological properties."
- “Global topological property is a basic factor in perceptual organization."



陈霖院士

Lin Chen. Topological structure in visual perception. Science, 1982, 218:699-700.

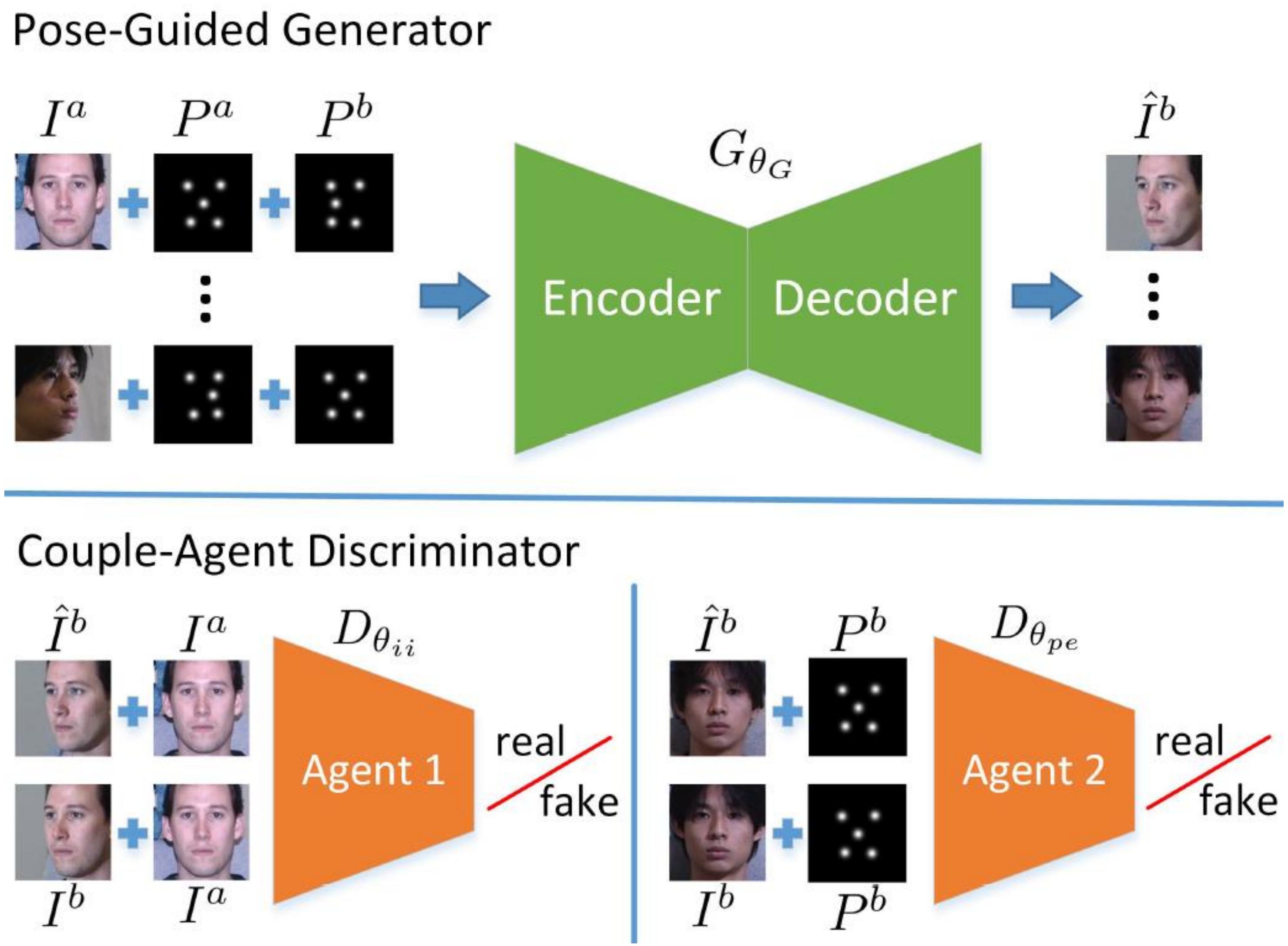
Framework — Couple-Agent Pose-Guided GAN



Contributions:

1. **Couple-Agent Pose-Guided GAN** (CAPG-GAN) is proposed for face rotation in 2D space.

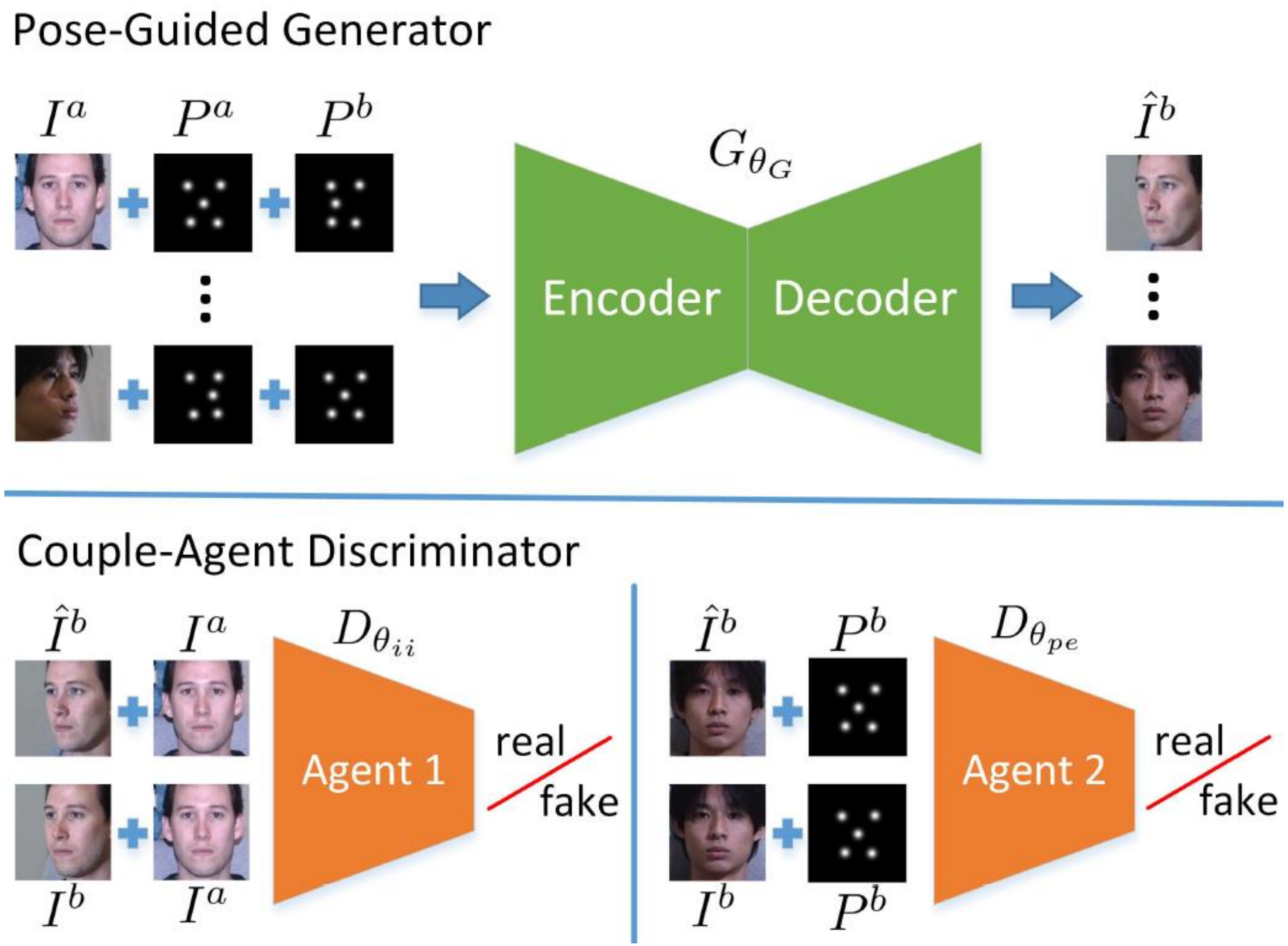
Framework — Couple-Agent Pose-Guided GAN



Contributions:

2. **Pose-guided generator** uses landmark heatmaps as controllable signals to synthesize arbitrary poses.

Framework — Couple-Agent Pose-Guided GAN

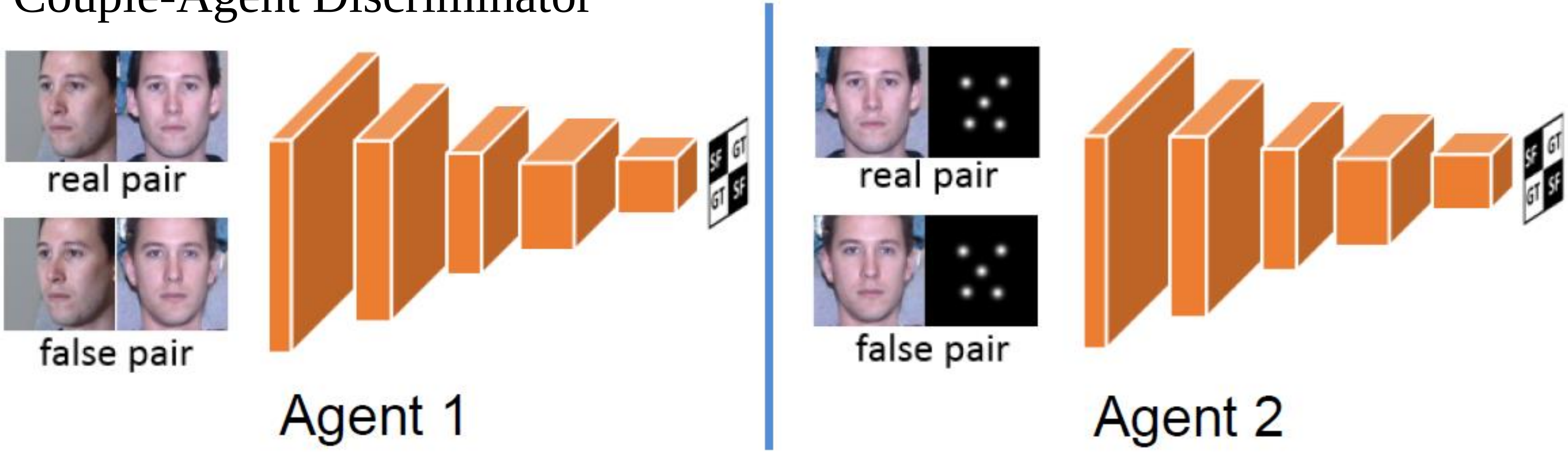


Contributions:

3. **Couple-agent discriminator** efficiently combines prior domain knowledge of poses and topological structure of faces to reinforce the realism.

Framework — Couple-Agent Pose-Guided GAN

Couple-Agent Discriminator



- **Rotation Pose:** Agent 1 distinguishes synthesized from natural face images, as well as the distinction of rotated poses.
- **Topological Structure:** Agent 2 discriminates the diversity of facial structure and captures the topological information.

Losses

- Conditional Adversarial Loss (Coupel-Agent Discriminator)

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

Fake Pair

Real Pair

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

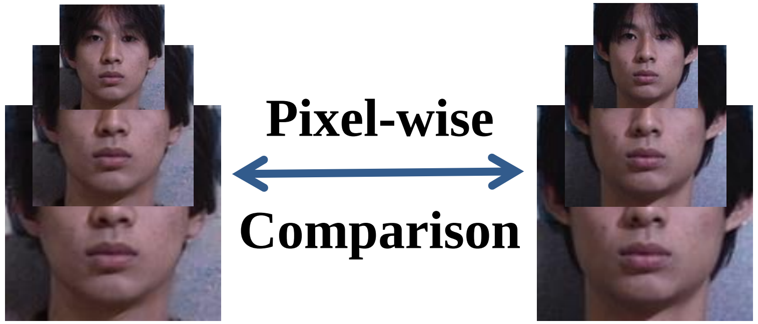
Fake Pair

Real Pair

Losses

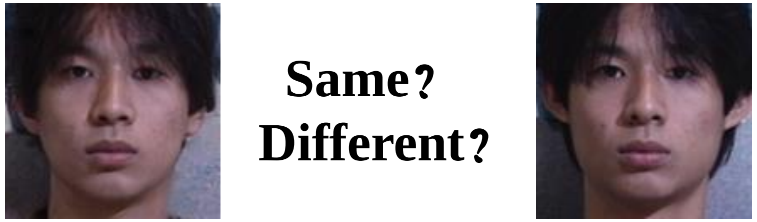
- Loss Multi-Scale Pixel-Wise Loss

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



- Identity Preserving Loss

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



- Total Variation Regularization

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

Results — Multi-PIE Frontalization

Figure 2. Synthesis results of different methods under the pose of 45° (first two rows) and 30° (last row).

[14] R. Huang, etal. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. In ICCV, 2017.
[28] L. Tran, etal. Disentangled representation learning gan for pose-invariant face recognition. In CVPR, 2017.
[34] J. Yim, etal. Rotating your face using multi-task deep neural network. In CVPR, 2015.
[7] A. Ghodrati, etal. Towards automatictowards automatic image editing: Learning to see another you. In BMVC, 2016.
[39] X. Zhu, etal. High-fidelity pose and expression normalization for face recognition in the wild. In CVPR, 2015.
[12] T. Hassner, etal. Effective face frontalization in unconstrained images. In CVPR, 2015.

Results — Multi-PIE Frontalization

Figure 3. Synthesis results of different methods under the pose of 75° (first two rows) and 90° (last two rows).

[14] R. Huang, etal. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. In ICCV, 2017.

[35] X. Yin, X. etal. Towards large-pose face frontalization in the wild. In ICCV, 2017.

Results — Multi-PIE Frontalization

Table 1. Rank-1 recognition rates (%) across views and illuminations under Setting 1.

Method	±90°	±75°	±60°	±45°	±30°	±15°
CPF[34]	-	-	-	71.65	81.05	89.45
Hassner *et al.* [12]	-	-	44.81	74.68	89.59	96.78
HPN[5]	29.82	47.57	61.24	72.77	78.26	84.23
FIP_40[40]	31.37	49.10	69.75	85.54	92.98	96.30
c-CNN Forest[32]	47.26	60.66	74.38	89.02	94.05	96.97
TP-GAN[14]	64.03	84.10	92.93	**98.58**	**99.85**	99.78
Light CNN[29]	9.00	32.35	73.30	97.45	99.80	99.78
CAPG-GAN	**77.10**	**87.40**	**93.74**	98.28	99.37	**99.95**

Table 2. Rank-1 recognition rates (%) across views, illuminations and sessions under Setting 2.

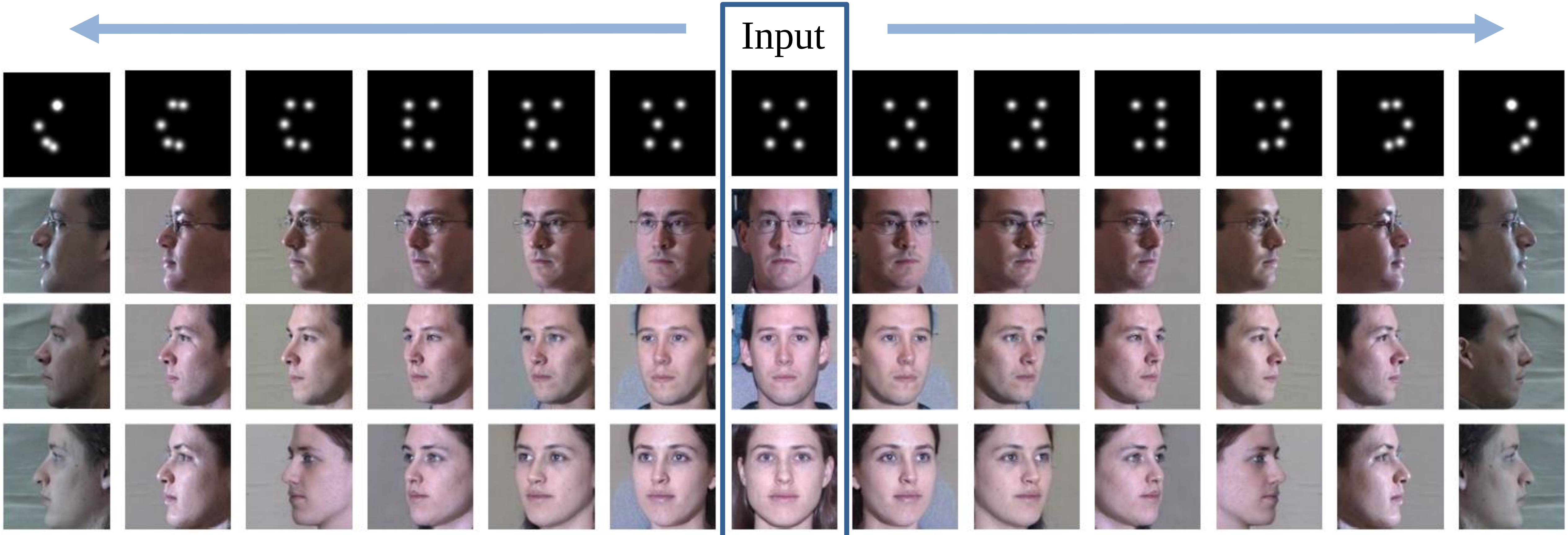
Method	±90°	±75°	±60°	±45°	±30°	±15°
FIP+LDA[40]	-	-	45.9	64.1	80.7	90.7
MVP+LDA[41]	-	-	60.1	72.9	83.7	92.8
CPF[34]	-	-	61.9	79.9	88.5	95.0
DR-GAN[28]	-	-	83.2	86.2	90.1	94.0
FF-GAN[35]	61.2	77.2	85.2	89.7	92.5	94.6
TP-GAN[14]	64.64	77.43	87.72	95.38	98.06	98.68
Light CNN[29]	5.51	24.18	62.09	92.13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

[5] C. Ding and D. Tao. Pose-invariant face recognition with homography-based normalization. PR, 66:144–152, 2017.
[12] T. Hassner, etal. Effective face frontalization in unconstrained images. In CVPR, 2015.
[14] R. Huang, etal. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. In ICCV, 2017.
[28] L. Tran, etal. Disentangled representation learning gan for pose-invariant face recognition. In CVPR, 2017.
[29] Wu X, He R, Sun Z, et al. A light CNN for deep face representation with noisy labels. TIFS, 2018.
[32] C. Xiong, X. Zhao, D. Tang, K. Jayashree, S. Yan, and T. K. Kim. Conditional convolutional neural network for modality-aware face recognition. In ICCV, 2015.
[34] J. Yim, etal. Rotating your face using multi-task deep neural network. In CVPR, 2015.
[35] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker. Towards large-pose face frontalization in the wild. In ICCV, 2017.
[40] Z. Zhu, P. Luo, X. Wang, and X. Tang. Deep learning identity-preserving face space. In ICCV, 2013.
[41] Z. Zhu, P. Luo, X. Wang, and X. Tang. Multi-view perceptron: a deep model for learning face identity and view representations. In NIPS, 2014.

Results — Multi-PIE Rotation

Figure 5. Synthesized frontal and profile results by CAPG-GAN under different poses. From top left to bottom right, the poses are 90°, 75°, 60°, 45°, 30° and 15°. The first and third images in each column are ground truth, and the second and forth images are synthesized.

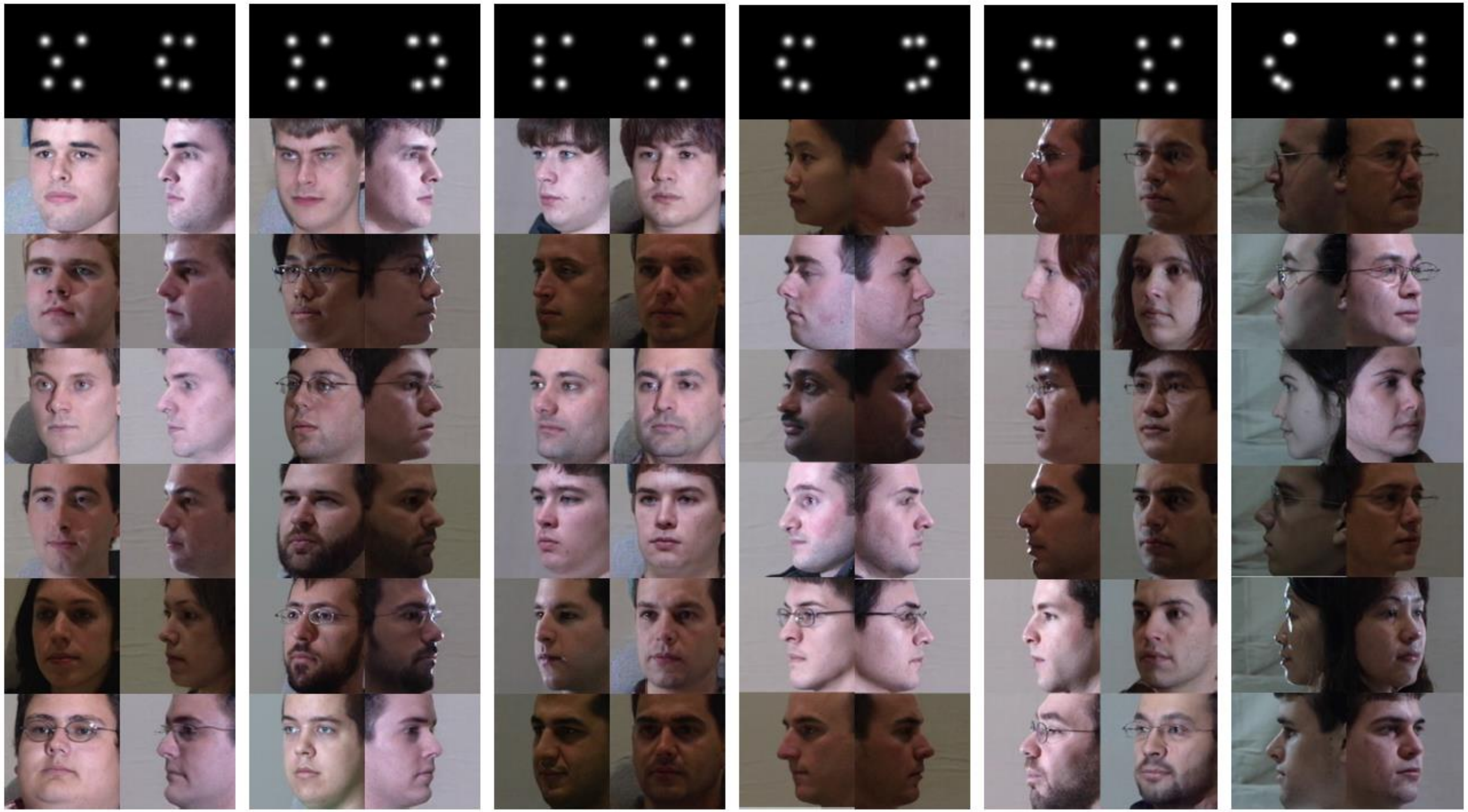
Results — Multi-PIE Rotation



(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°

Figure 6. Synthesis results of different target pose embeddings.

Results — Multi-PIE Rotation

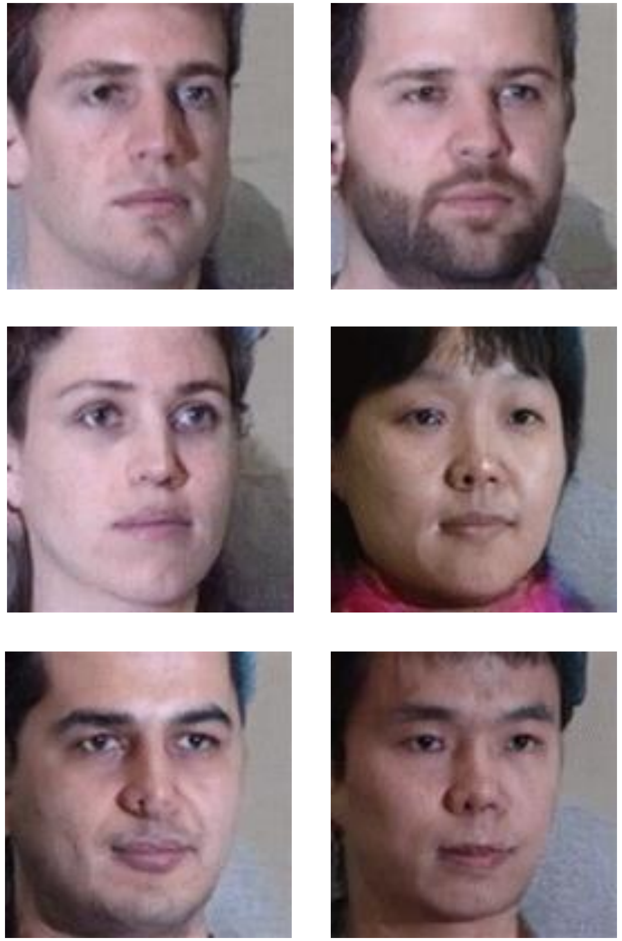


(a) $+15^\circ \rightarrow +60^\circ$ (b) $+30^\circ \rightarrow -60^\circ$ (c) $+45^\circ \rightarrow +15^\circ$ (d) $+60^\circ \rightarrow -75^\circ$ (e) $+75^\circ \rightarrow +30^\circ$ (f) $+90^\circ \rightarrow -45^\circ$

Figure 4. Synthesis results of various poses. For a pair of face images in one column, the left is the source image and the right is the synthesized one guided by the landmarks above.

Results — Multi-PIE Rotation

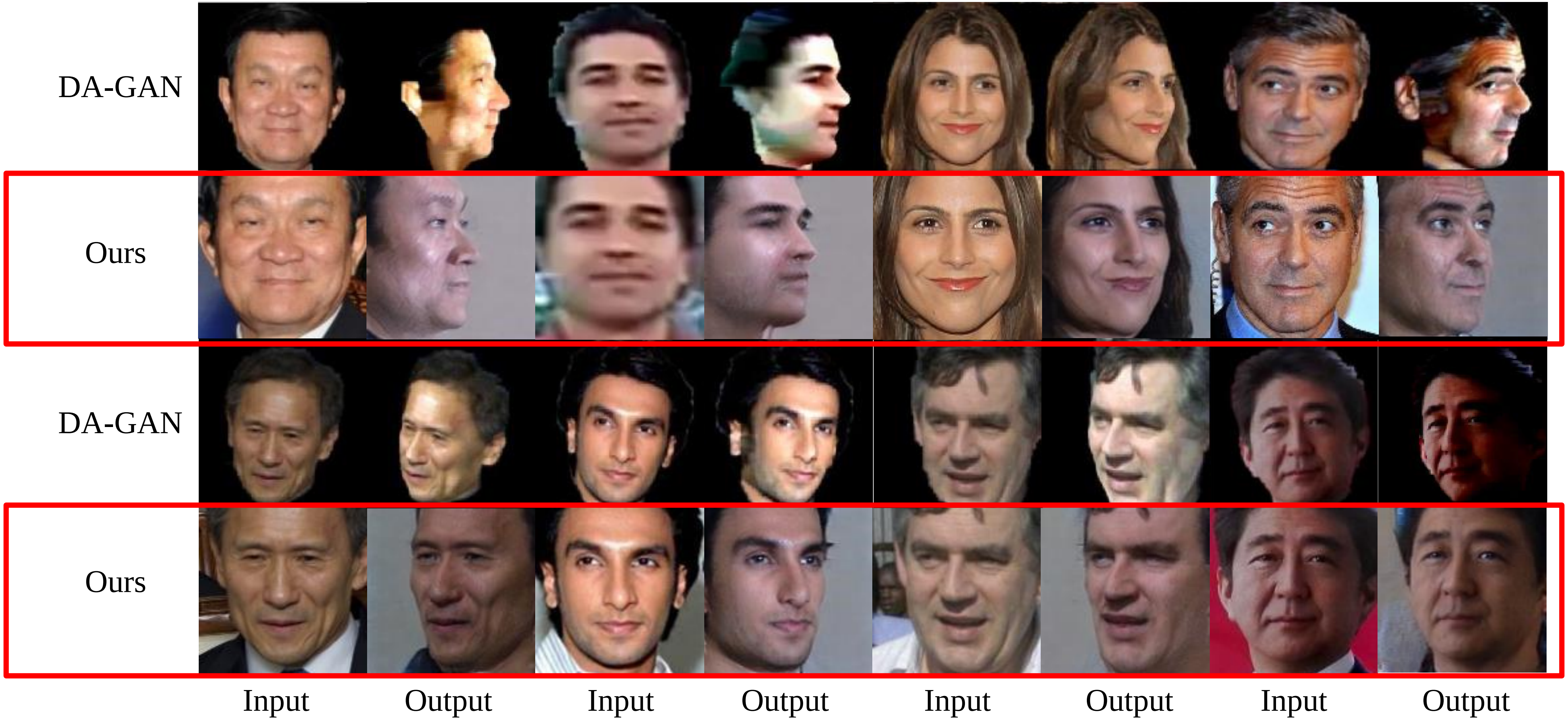
Figure 5. Synthesis results of landmark interpolation.



Results — LFW

TP-GAN: R. Huang, S. Zhang, T. Li, and R. He. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. In ICCV, 2017.
HPEN: X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li. High-fidelity pose and expression normalization for face recognition in the wild. In CVPR, 2015.
LFW-3D: T. Hassner, S. Harel, E. Paz, and R. Enbar. Effective face frontalization in unconstrained images. In CVPR, 2015.

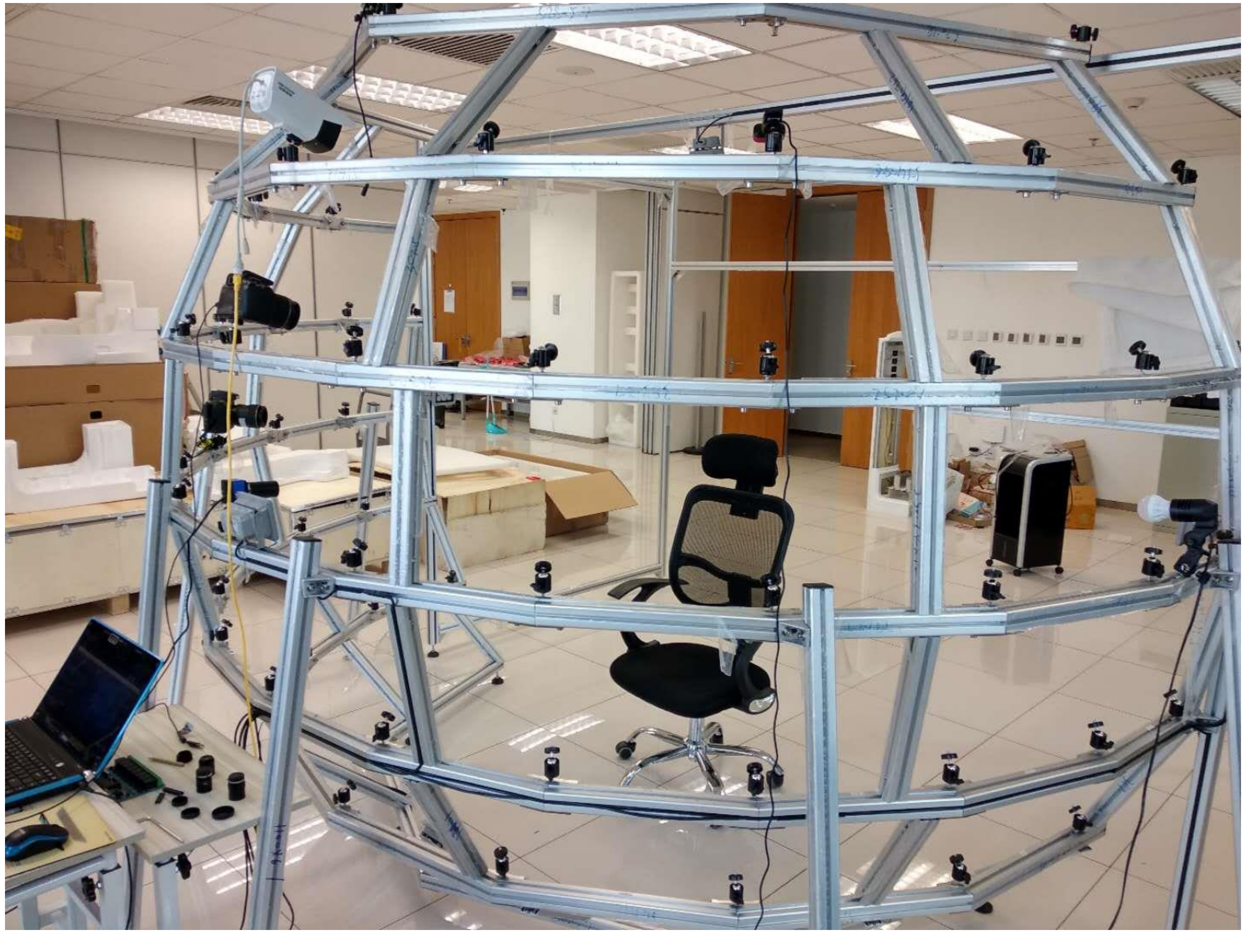
Results — IJBA

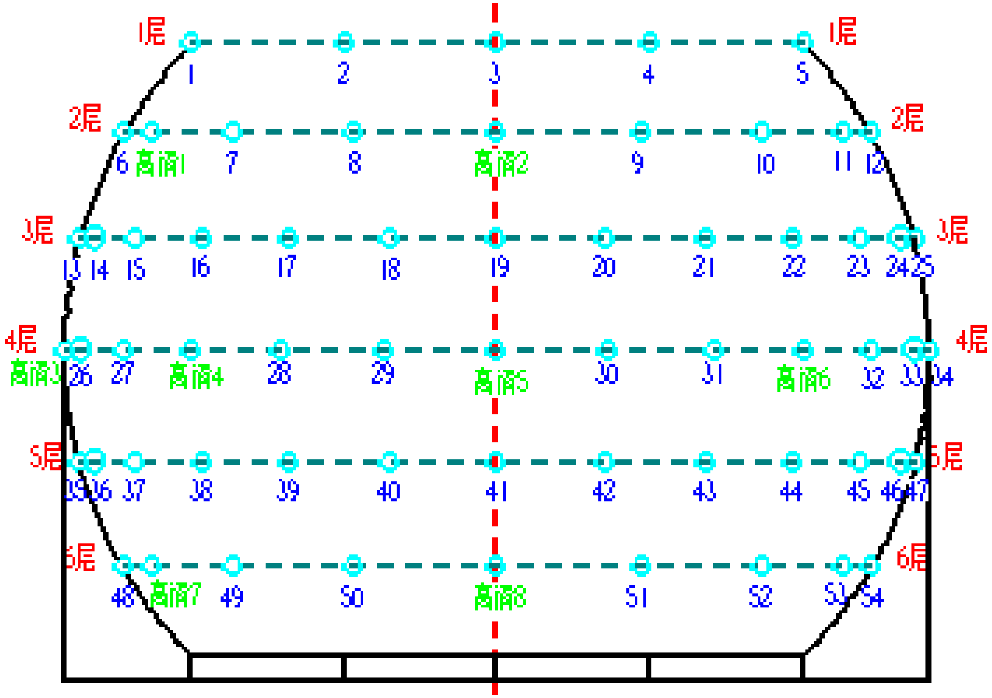


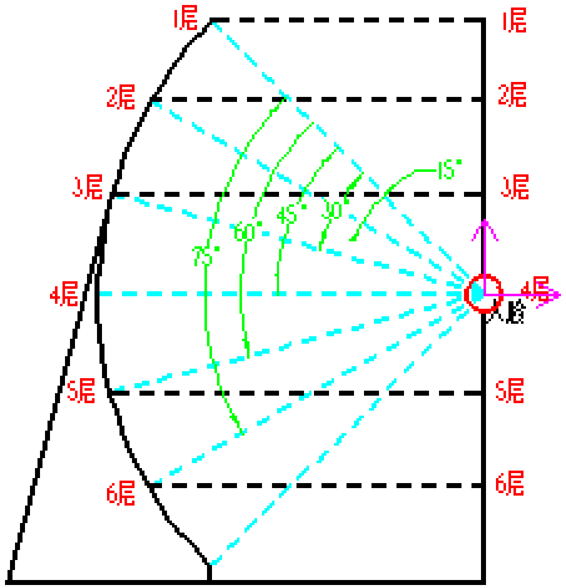
DA-GAN: J. Zhao, L. Xiong, K. Jayashree, J. Li, F. Zhao, Z. Wang, S. Pranata, S. Shen, S. Yan, and J. Feng. Dual-agent gans for photorealistic and identity preserving profile face synthesis. In NIPS, 2017.

Going Deeper

- Multi-View High-Resolution Collection







HUAWEI

Thank You


Center for Research on Intelligent Perception and Computing
Institute of Automation, Chinese Academy of Sciences