Supplementary Material for
Physically-Based Face Rendering for NIR-VIS Face Recognition

Anonymous Author(s)
Affiliation
Address
`email`

1 NIR-VIS Face Generation

We provide an algorithm (Alg. 1) that describes our steps of generating the NIR-VIS face pairs, namely a) the VIS reflectance acquisition, b) the NIR-VIS transformation and c) the NIR-VIS pairs rendering. Moreover, we visualize the NIR-VIS facial images generated for multiple identities in Fig. 1.

2 NIR-VIS Face Recognition

t-SNE visualization To better understand the advantage of the generated NIR-VIS facial images as well as the proposed IDentity-based Maximum Mean Discrepancy (ID-MMD) loss, we visualize the identity feature distributions of the Oulu-CASIA NIR-VIS [2] and the LAMP-HQ [5] dataset. Specifically, for each dataset, we randomly select 10 identities from the testing set. For each identity, we randomly selected 10 NIR images and 10 VIS images. We visualize the distribution of features derived by the baseline network "$LC - 29^{\dagger}(\mathcal{L}_{id})$" and the proposed one "$LC - 29^{\dagger}$ + Fake ($\mathcal{L}_{id}$ + $\mathcal{L}_{idmmd}$)" with t-SNE [4]. The visualization results are demonstrated in Fig. 2. Different identities are denoted by different colors. As can be seen, after involving the generated NIR-VIS facial images and the ID-MMD loss, the NIR-VIS features of the same identity are pulled closer. Meanwhile, for each identity, the features within the NIR/VIS domain distribute more compactly. Such visualization results suggest that the proposed method can effectively reduce both intra-modality and inter-modality discrepancies.

ROC curves The Receiver Operating Characteristic (ROC) curves of our method on four benchmark datasets, including CASIA NIR-VIS 2.0 [3], LAMP-HQ [5], Oulu-CASIA NIR-VIS [2], and BUAA-VisNir [1], are shown in Fig. 3. We compare the ROC curves of our models before and after the fine-tuning. As can be seen from the green curve, before involving any existing NIR-VIS face recognition datasets, our method has achieved comparable performance with the state-of-the-art methods in Table 4 of the main paper. After fine-tuning on the target NIR-VIS face recognition datasets, the performances are further boosted, as indicated by the red curves.

References

[1] Di Huang, Jia Sun, and Yunhong Wang. The buaa-visnir face database instructions. 2012.

[2] Jie Chen, D. Yi, Jimei Yang, Guoying Zhao, S. Z. Li, and M. Pietikainen. Learning mappings for face synthesis from near infrared to visual light images. In *CVPR*, 2009.

[3] Stan Li, Dong Yi, Zhen Lei, and Shengcai Liao. The casia nir-vis 2.0 face database. In *CVPR Workshops*, 2013.

[4] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *JMLR*, 2008.

[5] Aijing Yu, Haoxue Wu, Huaibo Huang, Zhen Lei, and Ran He. Lamp-hq: A large-scale multi-pose high-quality database and benchmark for nir-vis face recognition. *IJCV*, 2021.
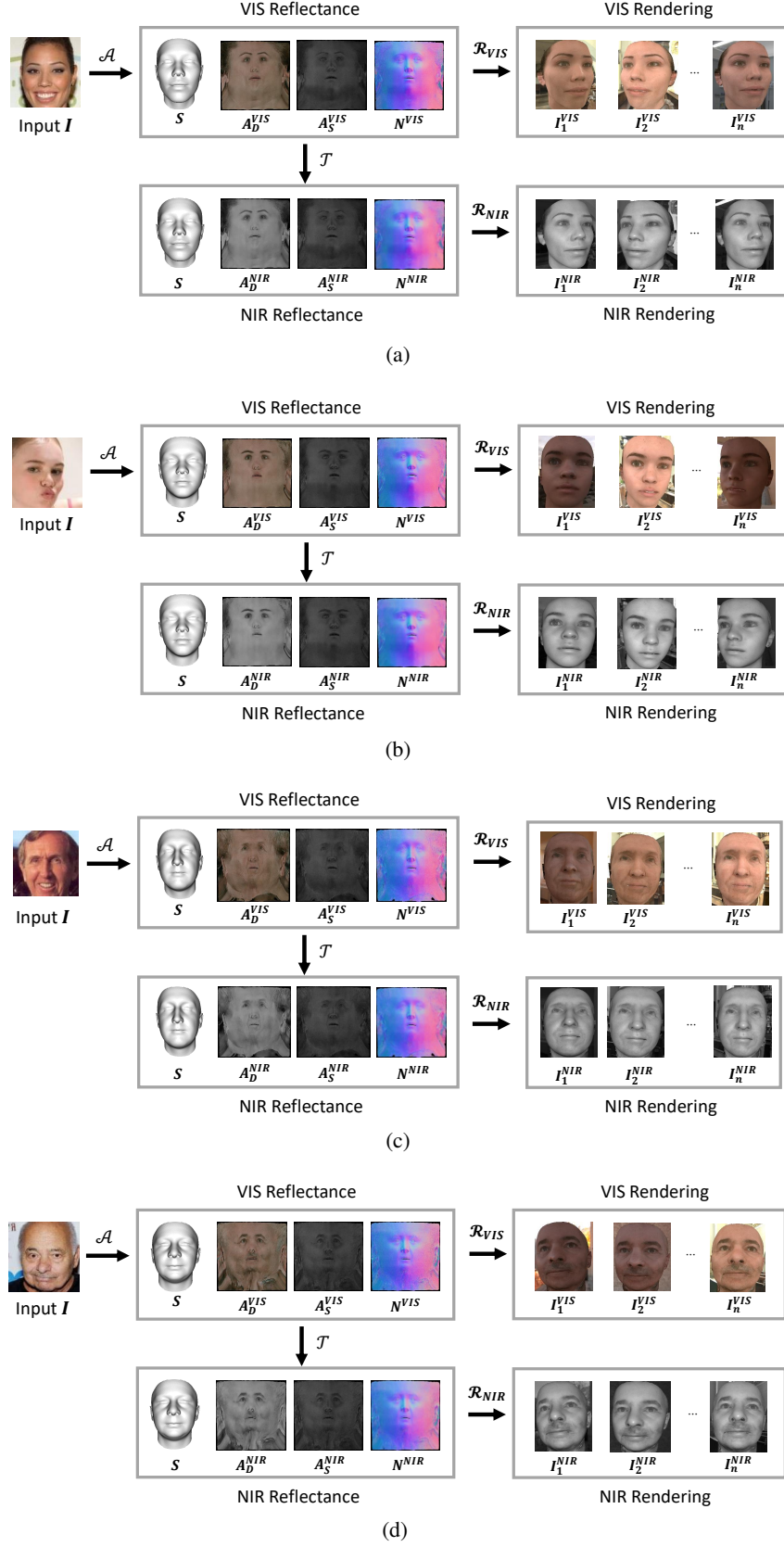
Figure 1: NIR-VIS facial image generation of four exemplar identities.

Algorithm 1 NIR-VIS Data Generation


Require: $\mathbf{I}, w^R, w^G, w^{NIR}, \mathbf{N_S^R}, \mathbf{n_{render}}, \mathbf{E_f}, \mathbf{N_S^G}$

  # Acquire VIS Reflectance from CelebA dataset:

  for $i \leftarrow 1 \dots n$ **do**

    $\mathbf{S_i}, \mathbf{A_{D_i}^{VIS}}, \mathbf{A_{S_i}^{VIS}}, \mathbf{N_{S_i}^{VIS}} \leftarrow \mathcal{A}(\mathcal{F}(\mathbf{I_i}))$

  end for

  # Calculate Transformation parameters using measured data:

  $\sigma \leftarrow \arg\min_\sigma \left| \mathbf{N_S^R} - \mathcal{G}(\mathbf{N_S^G}, \sigma) \right|_2$

  # Transform VIS Reflectance to NIR domain:

  for $i \leftarrow 1 \dots n$ **do**

    $\mathbf{A_{D_i}^{NIR}}, \mathbf{A_{S_i}^{NIR}}, \mathbf{N_{S_i}^{NIR}} = \mathcal{T}(\mathbf{A_{D_i}^{VIS}}, \mathbf{A_{S_i}^{VIS}}, \mathbf{N_{S_i}^{VIS}}, w^{NIR}, \sigma)$

  end for

  # Render NIR-VIS Pairs

  for $i \leftarrow 1 \dots n$ **do**

    for $j \leftarrow 1 \dots n_{render}$ **do**

      Sample $\mathbf{M_j}, \mathbf{E_j}$

      $\mathbf{E_j^{NIR}} = \frac{\mathbf{E_j} + \mathbf{E_f}}{||\mathbf{E_j} + \mathbf{E_f}||}$

      $\mathbf{R_i^{VIS}} \leftarrow \left[\mathbf{A_{D_i}^{VIS}}, \mathbf{A_{S_i}^{VIS}}, \mathbf{N_{S_i}^{VIS}} \right]$

      $\mathbf{R_i^{NIR}} \leftarrow \left[\mathbf{A_{D_i}^{NIR}}, \mathbf{A_{S_i}^{NIR}}, \mathbf{N_{S_i}^{NIR}} \right]$

      $\mathbf{I}_{i,j}^{VIS} = \mathcal{R}_i^{VIS}(\mathbf{S_i}, \mathbf{R_i^{VIS}}, \mathbf{M_j}, \mathbf{E_j})$

      $\mathbf{I}_{i,j}^{NIR} = \mathcal{R}_i^{NIR}(\mathbf{S_i}, \mathbf{R_i^{NIR}}, \mathbf{M_j}, \mathbf{E_j^{NIR}})$

    end for

  end for

  return $\mathbf{I}^{VIS}, \mathbf{I}^{NIR}$
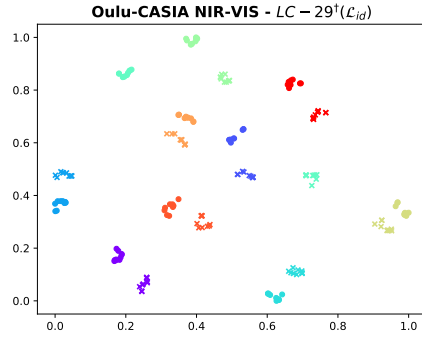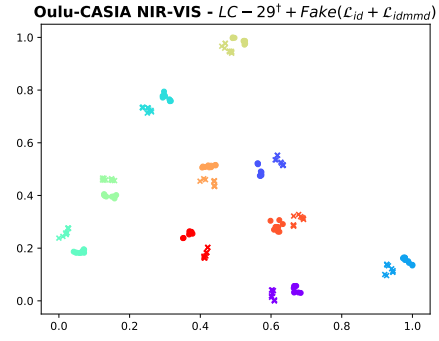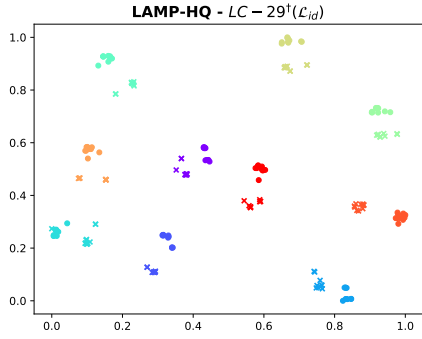
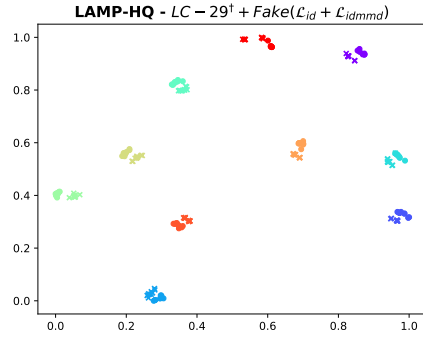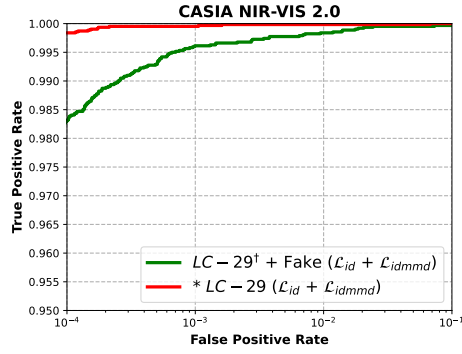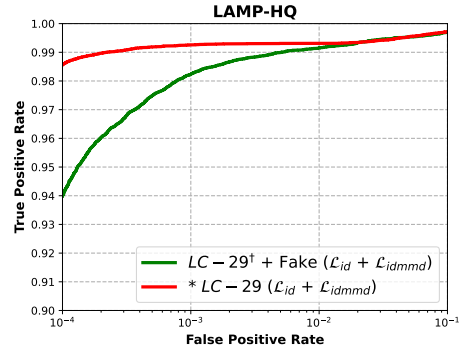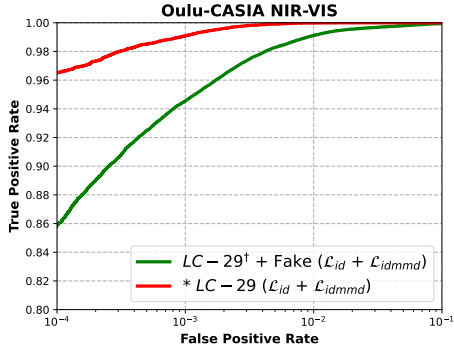
Figure 2: t-SNE [4] visualization of features of 10 identities randomly selected from (a)(b) Oulu-CASIA NIR-VIS, and (c)(d) LAMP-HQ. Different identities are denoted by different colors. ●: NIR images; ×: VIS images. (Best viewed in color).
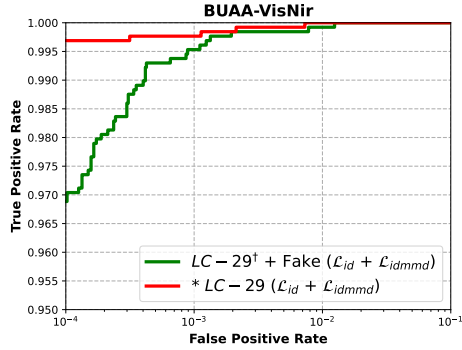
CASIA NIR-VIS 2.0

$LC - 29^{\dagger} +$ Fake $(\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

$* LC - 29 \; (\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

(a)

LAMP-HQ

$LC - 29^{\dagger} +$ Fake $(\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

$* LC - 29 \; (\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

(b)

Oulu-CASIA NIR-VIS

$LC - 29^{\dagger} +$ Fake $(\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

$* LC - 29 \; (\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

(c)

BUAA-VisNir

$LC - 29^{\dagger} +$ Fake $(\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

$* LC - 29 \; (\mathcal{L}_{id} + \mathcal{L}_{idmmd})$

(d)

Figure 3: ROC curves on (a) CASIA NIR-VIS 2.0, (b) LAMP-HQ, (c) Oulu-CASIA NIR-VIS, and (d) BUAA-VisNir.