APPENDICES OF HYPERSPHERE FACE UNCERTAINTY LEARNING

Anonymous authors
Paper under double-blind review

A PROOF OF THEOREM 1

Proof. By leveraging the asymptotic expansion of the modified Bessel function of the first kind (developed by Hermann Hankel), for any complex number z with large $|z|$ and $|\arg z| < \pi/2$,

$$\mathcal{I}_\alpha(z) \sim \frac{e^z}{\sqrt{2\pi z}}\left(1 - \frac{4\alpha^2-1}{8z} + \frac{\left(4\alpha^2-1\right)\left(4\alpha^2-9\right)}{2!(8z)^2} - \frac{\left(4\alpha^2-1\right)\left(4\alpha^2-9\right)\left(4\alpha^2-25\right)}{3!(8z)^3} + \cdots\right), \tag{1}$$

we have $\mathcal{I}_{d/2-1}(\kappa) \sim e^\kappa/\sqrt{2\pi\kappa}$ as $\kappa \to \infty$. Then, the r-radius vMF posterior can be rewritten as

$$\begin{aligned} p(\mathbf{z}|\mathbf{x}) &= \frac{\kappa^{d/2-1}}{\left(\sqrt{2\pi}r\right)^d \mathcal{I}_{d/2-1}(\kappa)} \exp\left(\frac{\kappa}{r}\boldsymbol{\mu}^T\mathbf{z}\right) \\ &= \frac{\sqrt{2\pi\kappa}\cdot\kappa^{d/2-1}}{\left(\sqrt{2\pi}r\right)^d \exp\left(\kappa\left(1-\frac{1}{r}\boldsymbol{\mu}^T\mathbf{z}\right)\right)} \end{aligned} \tag{2}$$

Therefore, when $\boldsymbol{\mu} = \mathbf{z}/r$, $p(\mathbf{z}|\mathbf{x}) \to \infty$ as $\kappa \to \infty$; otherwise, $p(\mathbf{z}|\mathbf{x}) \to 0$ as $\kappa \to \infty$. □

B PROOF OF COROLLARY 1

Proof. As $\kappa_\mathbf{x} \to \infty$, r-vMF$(\boldsymbol{\mu}_\mathbf{x}, \kappa_\mathbf{x}) \to \delta(\mathbf{z} - r\boldsymbol{\mu}_\mathbf{x})$. Then, $D_{\mathrm{KL}}(\delta(\mathbf{z}-r\boldsymbol{\mu}_\mathbf{x})||r\text{-vMF}(\boldsymbol{\mu}_\mathbf{x}, \kappa_\mathbf{x})) \to 0$. □

C FEATURE FUSION

In the cases where one subject has multiple face images (observations), it is desirable to build a compact distributional representation of this particular subject. In this section, we derive such distributional representation given multiple observations of the same identity and show that distributional parameters can be updated iteratively. Formally, given observations $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$, assuming the posterior distribution $p(\mathbf{z}|\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ is r-radius vMF with known parameters $\tilde{\boldsymbol{\mu}}_n$ and $\tilde{\kappa}_n$, we obtain the posterior given $(n+1)$ observations

$$p(\mathbf{z}|\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_{n+1}) = C \cdot \frac{p(\mathbf{z}|\mathbf{x}_{n+1})}{p(\mathbf{z})} p(\mathbf{z}|\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) \tag{3}$$

where C is a constant. Further assume that the prior $p(\mathbf{z})$ is uniform distribution defined in $r\mathbb{S}^{d-1}$. Then, it can be shown that

$$p(\mathbf{z}|\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, \mathbf{x}_{n+1}) \propto \exp\left(\frac{1}{r}(\kappa_{n+1}\boldsymbol{\mu}_{n+1} + \tilde{\kappa}_n\tilde{\boldsymbol{\mu}}_n)^T\mathbf{z}\right) := \exp\left(\frac{\tilde{\kappa}_{n+1}}{r}\tilde{\boldsymbol{\mu}}_{n+1}^T\mathbf{z}\right) \tag{4}$$

Therefore, we obtain iterative updating formulae for $\tilde{\boldsymbol{\mu}}_{n+1}$ and $\tilde{\kappa}_{n+1}$:

$$\tilde{\kappa}_{n+1} = ||\kappa_{n+1}\boldsymbol{\mu}_{n+1} + \tilde{\kappa}_n\tilde{\boldsymbol{\mu}}_n||_2, \quad \tilde{\boldsymbol{\mu}}_{n+1} = (\kappa_{n+1}\boldsymbol{\mu}_{n+1} + \tilde{\kappa}_n\tilde{\boldsymbol{\mu}}_n)/\tilde{\kappa}_{n+1}. \tag{5}$$

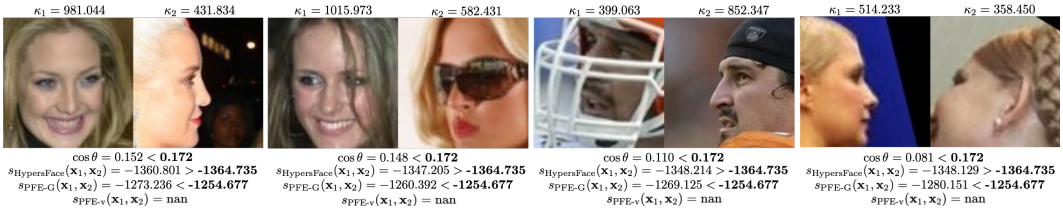

Figure 1: False negative examples made by PFE while being true positive by HypersFace, where $\cos\theta$ is the cosine distance of a verification pair $\mathbf{x}_1, \mathbf{x}_2$, $s(\cdot, \cdot)$ is mutual likelihood score and κ_1, κ_2 are the corresponding concentration values. Thresholds are set to -1254.677 and -1364.735 for PFE (accuracy: 88.210) and HypersFace (accuracy: 88.883), respectively, on the CPLFW benchmark.

D Mutual Likelihood Score for Hyperspherical Latents

We adopt mutual likelihood score proposed in (Shi & Jain, 2019) to conduct feature similarity comparison. The mutual likelihood score of two faces, $\mathbf{x}_i$ and $\mathbf{x}_j$, is defined as $s(\mathbf{x}_i, \mathbf{x}_j) = \log p(\mathbf{z}_i = \mathbf{z}_j)$. We show that a closed-form mutual likelihood score can be obtained for hyperspherical latents:

$$
\begin{aligned}
s(\mathbf{x}_i, \mathbf{x}_j) &= \log p(\mathbf{z}_i = \mathbf{z}_j) \\
&= \log \iint_{r\mathbb{S}^{d-1} \times r\mathbb{S}^{d-1}} p(\mathbf{z}_i|\mathbf{x}_i) p(\mathbf{z}_j|\mathbf{x}_j) \delta(\mathbf{z}_i - \mathbf{z}_j) d\mathbf{z}_i d\mathbf{z}_j \\
&= \log \frac{\mathcal{C}_d(\kappa_i)\mathcal{C}_d(\kappa_j)}{r^{2d}} \int_{r\mathbb{S}^{d-1}} \exp\left(\frac{1}{r}(\kappa_i \boldsymbol{\mu}_i + \kappa_j \boldsymbol{\mu}_j)^T \mathbf{z}\right) d\mathbf{z} \\
&= \log \frac{\mathcal{C}_d(\kappa_i)\mathcal{C}_d(\kappa_j)}{r^d \mathcal{C}_d(\tilde{\kappa})} \underbrace{\int_{r\mathbb{S}^{d-1}} \frac{\mathcal{C}_d(\tilde{\kappa})}{r^d} \exp\left(\frac{\tilde{\kappa}}{r} \tilde{\boldsymbol{\mu}}^T \mathbf{z}\right) d\mathbf{z}}_{=1} \\
&= \log \mathcal{C}_d(\kappa_i) + \log \mathcal{C}_d(\kappa_j) - \log \mathcal{C}_d(\tilde{\kappa}) - d \log r
\end{aligned}
\tag{6}
$$

where $\tilde{\kappa} = ||\mathbf{p}||_2$, $\mathbf{p} = (\kappa_i \boldsymbol{\mu}_i + \kappa_j \boldsymbol{\mu}_j)$, $\tilde{\boldsymbol{\mu}} = \mathbf{p}/||\mathbf{p}||_2$.

E Qualitative Analysis

We conduct qualitative analysis of how and why HypersFace outperforms PFE. Figure 2 shows some false positive examples made by PFE while being true negative examples made by HypersFace, and Figure 1 illustrates false negative examples made by PFE while being true positive examples made by HypersFace.

As shown in Figure 1, PFE fails to make correct predictions given genuine pairs due to the large pose variations (a)(d) and mask or sunglasses wearing (b)(c), whereas HypersFace is able to assign different concentration values to face images under different conditions. Higher concentration values indicate less uncertainty involved for our model to make predictions. Specifically, in the cases of large pose variations (a)(d) and severe or partial occlusions (b)(c), HypersFace adaptively gives lower concentration values to these samples, thereby making correct predictions that PFE does not.

Note that in Figure 1(d), HypersFace assigns relatively low concentration values to both samples but gives the first a slightly higher one than the second, since more facial clues can be seen from the first one. Figure 2 also demonstrates that PFE is unable to distinguish imposter pairs due to large pose variations (a)(c), partial occlusion (b) and image blur (d), whereas our model, HypersFace, successfully makes correct predictions by adaptively assigning proper concentration values to corresponding face images as in Figure 1.

References

Yichun Shi and Anil K Jain. Probabilistic face embeddings. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 6902–6911, 2019.

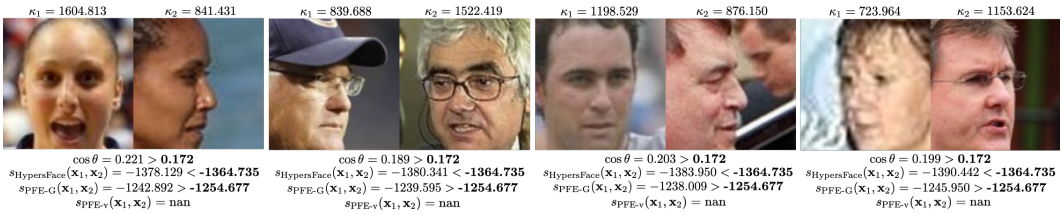

Figure 2: False positive examples made by PFE while being true positive by HypersFace, where $\cos\theta$ is the cosine distance of a verification pair $\mathbf{x}_1, \mathbf{x}_2$, $s(\cdot, \cdot)$ is mutual likelihood score and κ_1, κ_2 are the corresponding concentration values. Thresholds are set to -1254.677 and -1364.735 for PFE (accuracy: 88.210) and HypersFace (accuracy: 88.883), respectively, on the CPLFW benchmark.