A Characteristic Function-based Method for Bottom-up Human Pose Estimation (Supplementary Material)

Haoxuan Qu[1], Yujun Cai[2], Lin Geng Foo[1], Ajay Kumar[3], Jun Liu [1,*]
[1]Singapore University of Technology and Design, Singapore
[2]Nanyang Technological University, Singapore
[3]The Hong Kong Polytechnic University, Hong Kong
haoxuan_qu@mymail.sutd.edu.sg, yujun001@e.ntu.edu.sg, lingeng_foo@mymail.sutd.edu.sg
ajay.kumar@polyu.edu.hk, jun_liu@sutd.edu.sg

1. Additional Qualitative Results

In Fig. 1, we compare our method with the baseline method (HrHRNet-W32) [2], and present qualitative results on the COCO validation set. Note that for the baseline method (HrHRNet-W32) [2], we train the backbone model with the overall L2 loss, and for our method, we train the model with the loss $\hat{L}_{total}$ in Eq. 22 in the main paper. As shown, the baseline method can miss or misidentify body joints in some sub-regions of the predicted heatmap, whereas our method enables body joints in different sub-regions of the predicted heatmap to be located more accurately, which demonstrates the effectiveness of our method.

2. Additional Implementation Details about Loss

Recall that in the main paper, we construct $\hat{L}_k$ as:

$$\hat{L}_k = \sum_{m=1}^{M} \| \frac{\gamma}{2U} \big(\varphi^k_{D(H_p)}(\mathbf{t}_m) - \varphi^k_{D(H_g)}(\mathbf{t}_m) \big) \|_2^2 \tag{1}$$

where $\{\mathbf{t}_1, ..., \mathbf{t}_M\}$ denotes a set of M vectors randomly sampled from B_U. In this section, we introduce how we calculate $\hat{L}_k$ in more detail. Specifically, recall that $\varphi_D(\mathbf{t}) = E_{\mathbf{x}\sim D}[e^{i\langle \mathbf{t}, \mathbf{x}\rangle}]$ and $e^{i\langle \mathbf{t}, \mathbf{x}\rangle} = \cos(\langle \mathbf{t}, \mathbf{x}\rangle) + i \sin(\langle \mathbf{t}, \mathbf{x}\rangle)$. Then we can rewrite $\hat{L}_k$ as:

$$\begin{aligned}
\hat{L}_k =& \sum_{m=1}^{M} \| \frac{\gamma}{2U} \big(\varphi^k_{D(H_p)}(\mathbf{t}_m) - \varphi^k_{D(H_g)}(\mathbf{t}_m) \big) \|_2^2 \\
=& \frac{\gamma^2}{4U^2} \sum_{m=1}^{M} \| \varphi^k_{D(H_p)}(\mathbf{t}_m) - \varphi^k_{D(H_g)}(\mathbf{t}_m) \|_2^2 \\
=& \frac{\gamma^2}{4U^2} \sum_{m=1}^{M} \| E_{\mathbf{x}\sim D(H_p)}[\cos(\langle \mathbf{t}_m, \mathbf{x}\rangle)] + i E_{\mathbf{x}\sim D(H_p)}[\sin(\langle \mathbf{t}_m, \mathbf{x}\rangle)] - E_{\mathbf{x}\sim D(H_g)}[\cos(\langle \mathbf{t}_m, \mathbf{x}\rangle)] - i E_{\mathbf{x}\sim D(H_g)}[\sin(\langle \mathbf{t}_m, \mathbf{x}\rangle)] \|_2^2 \\
=& \frac{\gamma^2}{4U^2} \sum_{m=1}^{M} \Big(\big(E_{\mathbf{x}\sim D(H_p)}[\cos(\langle \mathbf{t}_m, \mathbf{x}\rangle)] - E_{\mathbf{x}\sim D(H_g)}[\cos(\langle \mathbf{t}_m, \mathbf{x}\rangle)] \big)^2 + \big(E_{\mathbf{x}\sim D(H_p)}[\sin(\langle \mathbf{t}_m, \mathbf{x}\rangle)] - E_{\mathbf{x}\sim D(H_g)}[\sin(\langle \mathbf{t}_m, \mathbf{x}\rangle)] \big)^2 \Big)
\end{aligned} \tag{2}$$

From Eq. 2, we can then calculate $\hat{L}_k$ simply through $\sin$ and $\cos$ operations.

3. Proof of Lemma 1 in the Main Paper

In this section, we discuss Lemma 1 in the main paper.

*Corresponding Author

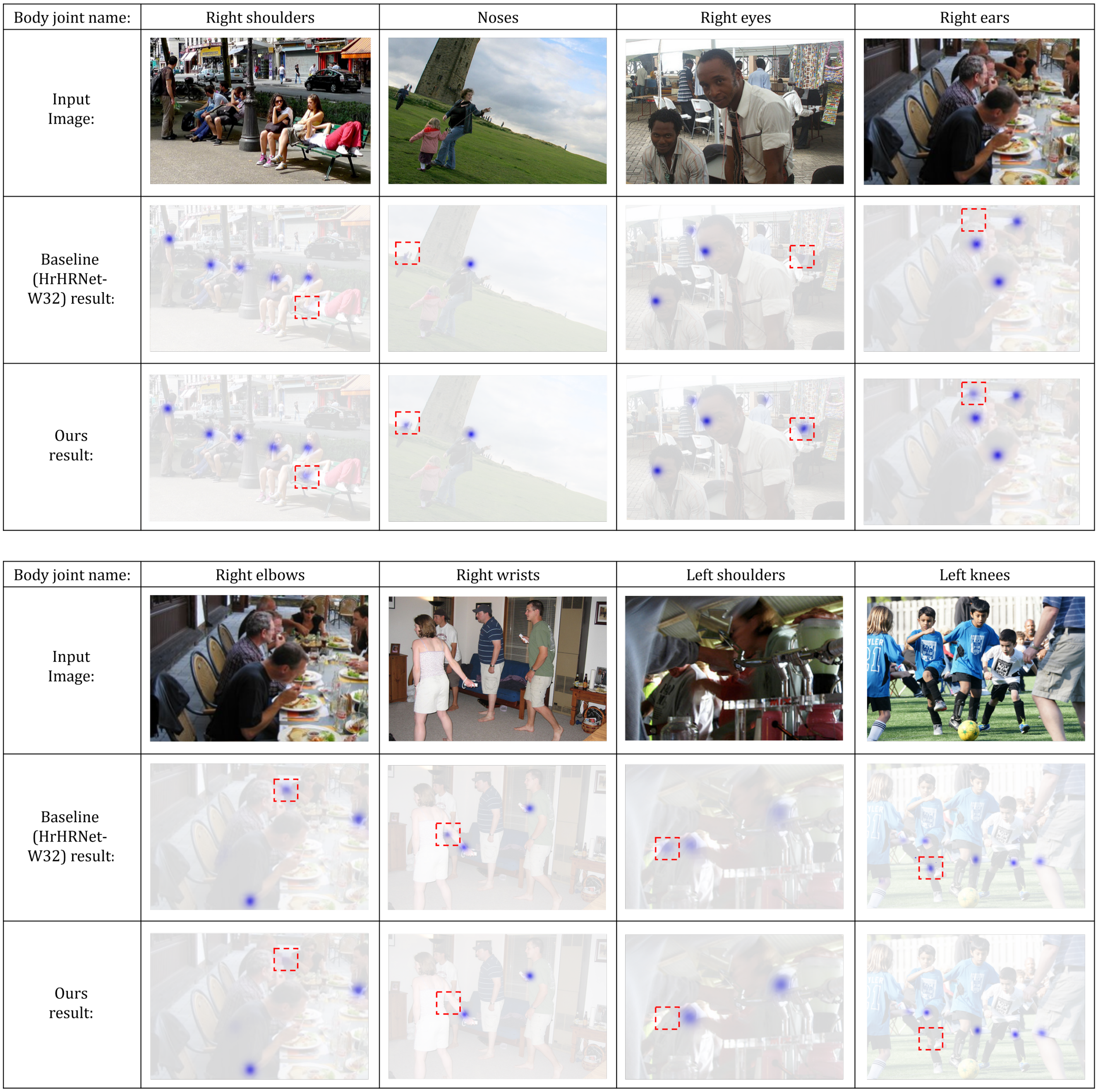

Figure 1. Qualitative results of our method and the baseline method [2].

Lemma 1. *Let φ_D be the characteristic function of a 2-dimensional distribution D. Let $R^r = [x_1^{lower}, x_1^{upper}] \times [x_2^{lower}, x_2^{upper}]$ a rectangular region, $R^e = \{x_1^{lower}, x_1^{upper}\} \times [x_2^{lower}, x_2^{upper}] \cup [x_1^{lower}, x_1^{upper}] \times \{x_2^{lower}, x_2^{upper}\}$ the edges of this region, and $R^v = \{x_1^{lower}, x_1^{upper}\} \times \{x_2^{lower}, x_2^{upper}\}$ the vertices of this region. Let $B_T = [-T, T] \times [-T, T]$. Denote $[D]_R$ the portion of the distribution D in R. $[D]_{R^r}$ can then be written as:*

$$[D]_{R^r} = \lim_{T\to\infty} \frac{1}{(2\pi)^2} \int_{B_T} \prod_{n=1}^{2} \left(\frac{e^{-it_n x_n^{lower}} - e^{-it_n x_n^{upper}}}{it_n}\right) \varphi_D(\boldsymbol{t})\, dt_1 dt_2 \; + \epsilon([D]_{R^r}) \tag{3}$$

where $\epsilon([D]_{R^r}) = \frac{[D]_{R^e}}{2} + \frac{[D]_{R^v}}{4}$ *and* $dt_1 dt_2$ *are calculated based on the Lebesgue measure.*

Proof. The proof sketch of Lemma 1 in the main paper is similar to the proof of Theorem 3.3.11 in [3]. Specifically, following the similar process of [3], we can first rewrite the right hand side of Eq. 3 as:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{(2\pi)^2} \int_{B_T} \prod_{n=1}^{2} \Big(\frac{e^{-it_n x_n^{lower}} - e^{-it_n x_n^{upper}}}{it_n}\Big) \varphi_D(\mathbf{t})\, dt_1 dt_2 \; + \epsilon([D]_{R^r}) \\
=& \lim_{T\to\infty} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \Big(\Big(\int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{lower}))}{t_1}\, dt_1 \; - \int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{upper}))}{t_1}\, dt_1 \; \Big) \\
& \qquad \times \Big(\int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{lower}))}{t_2}\, dt_2 \; - \int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{upper}))}{t_2}\, dt_2 \; \Big) \Big)\, dD \; + \epsilon([D]_{R^r})
\end{aligned} \tag{4}$$

Then following [3], we have the two equations below:

$$\lim_{T\to\infty} \Big(\int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{lower}))}{t_1}\, dt_1 \; - \int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{upper}))}{t_1}\, dt_1 \; \Big) = \begin{cases} 2\pi, & x_1^{lower} < x_1 < x_1^{upper} \\ \pi, & x_1 = x_1^{lower} \; or \; x_1 = x_1^{upper} \\ 0, & x_1 < x_1^{lower} \; or \; x_1 > x_1^{upper} \end{cases} \tag{5}$$

$$\lim_{T\to\infty} \Big(\int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{lower}))}{t_2}\, dt_2 \; - \int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{upper}))}{t_2}\, dt_2 \; \Big) = \begin{cases} 2\pi, & x_2^{lower} < x_2 < x_2^{upper} \\ \pi, & x_2 = x_2^{lower} \; or \; x_2 = x_2^{upper} \\ 0, & x_2 < x_2^{lower} \; or \; x_2 > x_2^{upper} \end{cases} \tag{6}$$

Then using Eq. 5 and Eq. 6, we have:

$$\begin{aligned}
&\lim_{T\to\infty} \Big(\Big(\int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{lower}))}{t_1}\, dt_1 \; - \int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{upper}))}{t_1}\, dt_1 \; \Big) \\
& \quad \times \Big(\int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{lower}))}{t_2}\, dt_2 \; - \int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{upper}))}{t_2}\, dt_2 \; \Big) \Big) \\
=& \begin{cases} (2\pi)^2, & (x_1^{lower}, x_1^{upper}) \times (x_2^{lower}, x_2^{upper}) \\ (2\pi) \times \pi, & \{x_1^{lower}, x_1^{upper}\} \times (x_2^{lower}, x_2^{upper}) \cup (x_1^{lower}, x_1^{upper}) \times \{x_2^{lower}, x_2^{upper}\} \\ \pi \times \pi, & \{x_1^{lower}, x_1^{upper}\} \times \{x_2^{lower}, x_2^{upper}\} \\ 0, & else \end{cases}
\end{aligned} \tag{7}$$

Then using Eq. 7, we can further rewrite Eq. 4 as:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{(2\pi)^2} \int_{B_T} \prod_{n=1}^{2} \Big(\frac{e^{-it_n x_n^{lower}} - e^{-it_n x_n^{upper}}}{it_n}\Big) \varphi_D(\mathbf{t})\, dt_1 dt_2 \; + \epsilon([D]_{R^r}) \\
=& \lim_{T\to\infty} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \Big(\Big(\int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{lower}))}{t_1}\, dt_1 \; - \int_{-T}^{T} \frac{\sin(t_1(x_1 - x_1^{upper}))}{t_1}\, dt_1 \; \Big) \\
& \qquad \times \Big(\int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{lower}))}{t_2}\, dt_2 \; - \int_{-T}^{T} \frac{\sin(t_2(x_2 - x_2^{upper}))}{t_2}\, dt_2 \; \Big) \Big)\, dD \; + \epsilon([D]_{R^r}) \\
=& ([D]_{R^r} - [D]_{R^e}) + \frac{[D]_{R^e} - [D]_{R^v}}{2} + \frac{[D]_{R^v}}{4} + \epsilon([D]_{R^r}) \\
=& [D]_{R^r} - \frac{[D]_{R^e}}{2} - \frac{[D]_{R^v}}{4} + \epsilon([D]_{R^r}) \\
=& [D]_{R^r}
\end{aligned} \tag{8}$$

□

4. Additional Details about Eq. 15 in the Main Paper

In this section, we discuss why Eq. 15 in the main paper holds. Specifically, as shown in [1], a mixture of Gaussian distributions with diagonal covariance matrices is a universal approximator of smooth distributions; some previous human pose estimation works [8, 9] also use a mixture of Gaussian distributions that effectively represents the predicted heatmaps. Therefore, we here rewrite $D(H_p)$ as a mixture of Gaussian distributions with diagonal covariance matrices. Besides, as H_g

is constructed via putting 2D Gaussian blobs centered at the GT coordinates of the body joints, we can also rewrite $D(H_g)$ as a mixture of Gaussian distributions with diagonal covariance matrices. After rewriting both $D(H_p)$ and $D(H_g)$ as a mixture of Gaussian distributions with diagonal covariance matrices, following [6], denoting $\mathbf{t} = (t_1, t_2)$, we can write $\varphi^k_{D(H_p)}(\mathbf{t})$ and $\varphi^k_{D(H_g)}(\mathbf{t})$ respectively as Eq. 9 and Eq. 10 below:

$$\begin{aligned}\varphi^k_{D(H_p)}(\mathbf{t}) &= \sum_{b=1}^{B} w_b e^{i\langle \mu_b, \mathbf{t}\rangle - \frac{1}{2}(\mathbf{t})^T \Sigma_b \mathbf{t}} \\ &= \sum_{b=1}^{B} w_b e^{i(\mu_{b,1}t_1 + \mu_{b,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{b,1})^2 + (t_2)^2(\sigma_{b,2})^2)} \\ &= \sum_{b=1}^{B} w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2 + (t_2)^2(\sigma_{b,2})^2)} \big(\cos(\mu_{b,1}t_1 + \mu_{b,2}t_2) + i \sin(\mu_{b,1}t_1 + \mu_{b,2}t_2) \big)\end{aligned} \tag{9}$$

where $\sum_{b=1}^{B} w_b = 1$, $w_b > 0$, $\mu_b = (\mu_{b,1}, \mu_{b,2})$ is the mean of the b-th Gaussian distribution component of $D(H_p)$, and $\Sigma_b = \begin{pmatrix} (\sigma_{b,1})^2 & 0 \\ 0 & (\sigma_{b,2})^2 \end{pmatrix}$ is the variance of the b-th Gaussian distribution component of $D(H_p)$.

$$\begin{aligned}\varphi^k_{D(H_g)}(\mathbf{t}) &= \sum_{c=1}^{C} w_c e^{i\langle \mu_c, \mathbf{t}\rangle - \frac{1}{2}(\mathbf{t})^T \Sigma_c \mathbf{t}} \\ &= \sum_{c=1}^{C} w_c e^{i(\mu_{c,1}t_1 + \mu_{c,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{c,1})^2 + (t_2)^2(\sigma_{c,2})^2)} \\ &= \sum_{c=1}^{C} w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2 + (t_2)^2(\sigma_{c,2})^2)} \big(\cos(\mu_{c,1}t_1 + \mu_{c,2}t_2) + i \sin(\mu_{c,1}t_1 + \mu_{c,2}t_2) \big)\end{aligned} \tag{10}$$

where $\sum_{c=1}^{C} w_c = 1$, $w_c > 0$, $\mu_c = (\mu_{c,1}, \mu_{c,2})$ is the mean of the c-th Gaussian distribution component of $D(H_g)$, and $\Sigma_c = \begin{pmatrix} (\sigma_{c,1})^2 & 0 \\ 0 & (\sigma_{c,2})^2 \end{pmatrix}$ is the variance of the c-th Gaussian distribution component of $D(H_g)$.

After that, further denoting $\mathbf{x} = (x_1, x_2)$, we have:

$$\begin{aligned}&\frac{\varphi^k_{D(H_p)}(\mathbf{t}) - \varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t}, \mathbf{x}\rangle} \\ =&\frac{e^{-i\langle \mathbf{t}, \mathbf{x}\rangle}}{(2\pi)^2} \Big(\sum_{b=1}^{B} w_b e^{i(\mu_{b,1}t_1 + \mu_{b,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{b,1})^2 + (t_2)^2(\sigma_{b,2})^2)} - \sum_{c=1}^{C} w_c e^{i(\mu_{c,1}t_1 + \mu_{c,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{c,1})^2 + (t_2)^2(\sigma_{c,2})^2)} \Big) \\ =&\frac{e^{-it_1x_1 - it_2x_2}}{(2\pi)^2} \Big(\sum_{b=1}^{B} w_b e^{i(\mu_{b,1}t_1 + \mu_{b,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{b,1})^2 + (t_2)^2(\sigma_{b,2})^2)} - \sum_{c=1}^{C} w_c e^{i(\mu_{c,1}t_1 + \mu_{c,2}t_2) - \frac{1}{2}((t_1)^2(\sigma_{c,1})^2 + (t_2)^2(\sigma_{c,2})^2)} \Big) \\ =&\frac{1}{(2\pi)^2} \Big(\sum_{b=1}^{B} w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2 + (t_2)^2(\sigma_{b,2})^2)} \big(\cos((\mu_{b,1} - x_1)t_1 + (\mu_{b,2} - x_2)t_2) + i \sin((\mu_{b,1} - x_1)t_1 + (\mu_{b,2} - x_2)t_2) \big) \\ &\quad - \sum_{c=1}^{C} w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2 + (t_2)^2(\sigma_{c,2})^2)} \big(\cos((\mu_{c,1} - x_1)t_1 + (\mu_{c,2} - x_2)t_2) + i \sin((\mu_{c,1} - x_1)t_1 + (\mu_{c,2} - x_2)t_2) \big) \Big) \\ =&Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t}) - \varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t}, \mathbf{x}\rangle} \Big) + i \times Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t}) - \varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t}, \mathbf{x}\rangle} \Big)\end{aligned} \tag{11}$$

where:

$$
\begin{aligned}
&Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\cos((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\cos((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
&Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\sin((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\sin((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)
\end{aligned}
\tag{13}
$$

Then denoting $\sigma=\min(\min_{b\in B}\sigma_{b,1},\min_{b\in B}\sigma_{b,2},\min_{c\in C}\sigma_{c,1},\min_{c\in C}\sigma_{c,2})$, we can find the upper bound and the lower bound of $Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)$ and $Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)$ respectively as:

$$
\begin{aligned}
&Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\cos((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\cos((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)\\
\le&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b-\sum_{c=1}^{C}(-w_c)\big)\\
=&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
&Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\cos((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\cos((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)\\
\ge&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{(2\pi)^2}\big(\sum_{b=1}^{B}(-w_b)-\sum_{c=1}^{C}w_c\big)\\
=&-\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}
\end{aligned}
\tag{15}
$$

$$
\begin{aligned}
&Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\sin((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\sin((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)\\
\leq&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b-\sum_{c=1}^{C}(-w_c)\big)\\
=&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
&Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\\
=&\frac{1}{(2\pi)^2}\big(\sum_{b=1}^{B}w_b e^{-\frac{1}{2}((t_1)^2(\sigma_{b,1})^2+(t_2)^2(\sigma_{b,2})^2)}\sin((\mu_{b,1}-x_1)t_1+(\mu_{b,2}-x_2)t_2)\\
&\qquad-\sum_{c=1}^{C}w_c e^{-\frac{1}{2}((t_1)^2(\sigma_{c,1})^2+(t_2)^2(\sigma_{c,2})^2)}\sin((\mu_{c,1}-x_1)t_1+(\mu_{c,2}-x_2)t_2)\big)\\
\geq&\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{(2\pi)^2}\big(\sum_{b=1}^{B}(-w_b)-\sum_{c=1}^{C}w_c\big)\\
=&-\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}
\end{aligned}
\tag{17}
$$

Then with Eq. 14 in the main paper rewritten as:

$$
\begin{aligned}
&\|\lim_{T\to\infty}\int_{B_T}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\,\|_2^2\\
=&\|\int_{B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\;+\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\,\|_2^2
\end{aligned}
\tag{18}
$$

we can rewrite the upper bound and the lower bound of the real part and the imaginary part of the latter term

$\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\, d\mathbf{x}\ \ d\mathbf{t}$ as:

$$\begin{aligned}
&Re\Big(\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\, d\mathbf{x}\ \ d\mathbf{t}\Big)\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\Big)\, d\mathbf{x}\ \ d\mathbf{t}\\
\leq&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{x}\ \ d\mathbf{t}\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U} \lambda(R^r_{sub})\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{t}\\
\leq&\frac{\lambda(R^r_{sub})e^{-\frac{1}{2}U^2\sigma^2}}{\pi\sigma^2}
\end{aligned} \tag{19}$$

$$\begin{aligned}
&Re\Big(\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\, d\mathbf{x}\ \ d\mathbf{t}\Big)\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} Re\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\Big)\, d\mathbf{x}\ \ d\mathbf{t}\\
\geq&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} -\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{x}\ \ d\mathbf{t}\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U} -\lambda(R^r_{sub})\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{t}\\
\geq&-\frac{\lambda(R^r_{sub})e^{-\frac{1}{2}U^2\sigma^2}}{\pi\sigma^2}
\end{aligned} \tag{20}$$

$$\begin{aligned}
&Im\Big(\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\, d\mathbf{x}\ \ d\mathbf{t}\Big)\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2} e^{-i\langle \mathbf{t},\mathbf{x}\rangle}\Big)\, d\mathbf{x}\ \ d\mathbf{t}\\
\leq&\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}} \frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{x}\ \ d\mathbf{t}\\
=&\lim_{T\to\infty}\int_{B_T\setminus B_U} \lambda(R^r_{sub})\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\, d\mathbf{t}\\
\leq&\frac{\lambda(R^r_{sub})e^{-\frac{1}{2}U^2\sigma^2}}{\pi\sigma^2}
\end{aligned} \tag{21}$$

$$
\begin{aligned}
&Im\Big(\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\Big)\\
&=\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}Im\Big(\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\Big)\,d\mathbf{x}\;d\mathbf{t}\\
&\geq\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}-\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\,d\mathbf{x}\;d\mathbf{t}\\
&=\lim_{T\to\infty}\int_{B_T\setminus B_U}-\lambda(R^r_{sub})\frac{e^{-\frac{1}{2}((t_1)^2\sigma^2+(t_2)^2\sigma^2)}}{2\pi^2}\,d\mathbf{t}\\
&\geq-\frac{\lambda(R^r_{sub})e^{-\frac{1}{2}U^2\sigma^2}}{\pi\sigma^2}
\end{aligned}
\tag{22}
$$

Then as shown from Eq. 19 to Eq. 22, roughly speaking, the values of both the real part and the imaginary part of $\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}$ decay exponentially when U becomes larger (as both their upper bound and their lower bound decay exponentially when U becomes larger). Thus, both the real part and the imaginary part of the latter term $\lim_{T\to\infty}\int_{B_T\setminus B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}$ outside a certain finite range B_U would be close to zero. Therefore, we can further approximate Eq. 14 in the main paper as:

$$
\|\lim_{T\to\infty}\int_{B_T}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\,\|_2^2\approx\|\int_{B_U}\int_{R^r_{sub}}\frac{\varphi^k_{D(H_p)}(\mathbf{t})-\varphi^k_{D(H_g)}(\mathbf{t})}{(2\pi)^2}e^{-i\langle\mathbf{t},\mathbf{x}\rangle}\,d\mathbf{x}\;d\mathbf{t}\,\|_2^2 \tag{23}
$$

5. Experiments on Top-down Methods

While our method is primarily designed for bottom-up human pose estimation to optimize the body joints over sub-regions of the predicted heatmap at the same time, we here also test the effectiveness of our method on top-down human pose estimation methods. Specifically, we here apply our method on various top-down methods including Simple Baseline [10], HRNet [7], and HRFormer [11]. We set input size to 384×288 and following [7, 10, 11], we report the following metrics: AP, AP^{50}, AP^{75}, AP^{M}, AP^{L}, and AR. As shown in Tab. 1, after incorporating our method in various top-down methods, we also observe a consistent performance improvement. A possible reason for this is that our method enables the background sub-regions of the predicted heatmap and the sub-regions of the predicted heatmap around the body joint to be optimized simultaneously, which thus is also helpful for top-down human pose estimation.

Table 1. We also compare with top-down methods on the **COCO val2017** set and the **COCO test-dev2017** set. Though our method is primarily designed for bottom-up human pose estimation, we observe that it is also helpful for top-down methods.

		COCO val2017						COCO test-dev2017					
Method	Backbone	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR
Simple Baseline [10]	ResNet-152	75.0	90.8	82.1	67.8	78.3	80.0	73.8	91.7	81.2	70.3	80.0	79.1
+ Ours	ResNet-152	76.1	91.7	83.1	69.1	79.4	80.8	74.9	92.5	82.8	71.3	80.7	80.0
HRNet [7]	HRNet-W32	76.7	91.9	83.6	73.2	83.2	81.6	74.9	92.5	82.8	71.3	80.9	80.1
+ Ours	HRNet-W32	77.6	92.2	84.3	73.5	84.2	82.0	75.9	92.8	83.7	72.6	81.6	81.0
HRNet [7]	HRNet-W48	77.1	91.8	83.8	73.5	83.5	81.8	75.5	92.5	83.3	71.9	81.5	80.5
+ Ours	HRNet-W48	78.1	92.4	84.8	74.1	84.8	82.4	76.5	93.0	84.3	73.1	82.0	81.5
HRFormer [11]	HRFormer-Base	78.0	92.2	84.8	74.3	84.6	82.6	76.2	92.7	83.8	72.5	82.3	81.2
+ Ours	HRFormer-Base	78.7	92.4	85.3	74.9	85.2	83.1	76.9	93.1	84.6	73.5	82.6	81.9

6. Dataset Licenses

The COCO dataset [5] is licensed under the following the Creative Commons Attribution 4.0 License. The CrowdPose dataset [4] is released for academic research only and it is free to researchers from educational or research institutes for non-commercial purposes.

References

[1] Samy Bengio. An introduction to statistical machine learning–em for gmms. *Dalle Molle Institute for Perceptual Artificial Intelligence (IDIAP) slides*, 2004. 3

[2] Bowen Cheng, Bin Xiao, Jingdong Wang, Honghui Shi, Thomas S Huang, and Lei Zhang. Higherhrnet: Scale-aware representation learning for bottom-up human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5386–5395, 2020. 1, 2

[3] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019. 3

[4] Jiefeng Li, Can Wang, Hao Zhu, Yihuan Mao, Hao-Shu Fang, and Cewu Lu. Crowdpose: Efficient crowded scenes pose estimation and a new benchmark. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10863–10872, 2019. 8

[5] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014. 8

[6] Weibo Shu, Jia Wan, Kay Chen Tan, Sam Kwong, and Antoni B Chan. Crowd counting in the frequency domain. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19618–19627, 2022. 4

[7] Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang. Deep high-resolution representation learning for human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5693–5703, 2019. 8

[8] Ben Usman, Andrea Tagliasacchi, Kate Saenko, and Avneesh Sud. Metapose: Fast 3d pose from multiple views without 3d supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6759–6770, 2022. 3

[9] Chen Wang, Feng Zhang, Xiatian Zhu, and Shuzhi Sam Ge. Low-resolution human pose estimation. *Pattern Recognition*, 126:108579, 2022. 3

[10] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking. In *Proceedings of the European conference on computer vision (ECCV)*, pages 466–481, 2018. 8

[11] Yuhui Yuan, Rao Fu, Lang Huang, Weihong Lin, Chao Zhang, Xilin Chen, and Jingdong Wang. Hrformer: High-resolution transformer for dense prediction. 2021. 8