Probabilistic Triangulation for Uncalibrated Multi-View 3D Human Pose Estimation

Boyuan Jiang[1,2], Lei Hu[1,2] Shihong Xia[1,2*]
[1]Institute of Computing Technology, Chinese Academy of Sciences;
[2]University of Chinese Academy of Sciences
{jiangboyuan20s, hulei19z, xsh}@ict.ac.cn

1. Formula details

1.1. heatmap-based representation

We define the uncalibrated 3D human pose estimation problem in the main text:

$$\min_{X_{3D},y} \quad E\left[proj(X^{3D},y)-X^{2D}\right] \tag{1}$$

When using heatmap-based representation of human pose, it is assumed that X^{2D} is the edge distribution of X^{3D}, and problem1 is transformed into:

$$\min_{y,X_{3D}} \quad \sum_i^N \int_{\sigma_i} \left\| x_i^{2D}(\sigma_i) - \int_{proj(z,y_i)=\sigma_i} x^{3D}(z)\,\mathrm{d}z \right\|_2 \mathrm{d}\sigma_i \tag{2}$$

where z denotes the body element in 3D space, σ_i denotes the area element of the ith camera plane, and $proj(\cdot)$ denotes the projection function.

In order to make the problem2 solvable, X^{3D} is transformed into an expression for X^{2D} and y, a process in which we use approximations:

$$\begin{aligned} x^{3D}(z) &= \frac{1}{N} \sum_{\substack{i \\ proj(z,y_i)=\sigma_i}}^N x_i^{2D}(\sigma_i) \\ x_i^{2D}(\sigma_i) &= \max_{proj(z,y_i)=\sigma_i} x^{3D}(z) \end{aligned} \tag{3}$$

Bringing in the equation 3 the detailed derivation proceeds as follows:

$$\begin{aligned} &\min_y \sum_i^N \int_{\sigma_i} \left\| x_i^{2D}(\sigma_i) - \int_{proj(z,y_i)=\sigma_i} x^{3D}(z)\,\mathrm{d}z \right\|_2 \mathrm{d}\sigma_i \\ \Leftrightarrow &\min_y \sum_i^N \left(\sum_{\sigma_i} \left\| x_i^{2D}(\sigma_i) - \frac{\sum_{z\in\{z|proj(z,y_i)=\sigma_i\}} (x^{3D}(z))^2}{\sum_z x^{3D}(z)} \right\|_2 x_i^{2D}(\sigma_i) \right) \\ &\quad / \left(\sum_{\sigma_i} x_i^{2D}(\sigma_i) \right) \\ \Leftrightarrow &\min_y \sum_i^N \sum_{\sigma_i} \left\| \sum_{z\in\{z|proj(z,y_i)=\sigma_i\}} x^{3D}(z) \left(x_i^{2D}(\sigma_i) - x^{3D}(z) \right) \right\|_2 x_i^{2D}(\sigma_i) \\ \Leftrightarrow &\min_y \sum_i^N \sum_{\sigma_i} \sum_{z\in\{z|proj(z,y_i)=\sigma_i\}} \left\| \left(x^{3D}(z) x_i^{2D}(\sigma_i) \right)^2 \left(x_i^{2D}(\sigma_i) - x^{3D}(z) \right) \right\|_2 \\ \Leftrightarrow &\min_y \sum_i^N \sum_z \left\| \left(x^{3D}(z) x_i^{2D}(\sigma_i) \right)^2 \left(x_i^{2D}(\sigma_i) - x^{3D}(z) \right) \right\|_2 \\ \Leftrightarrow &\min_y \sum_i^N \left\| \sum_z \left(x^{3D}(z) x_i^{2D}(\sigma_i) \right)^2 \left(x_i^{2D}(\sigma_i) - x^{3D}(z) \right) \right\|_2 \end{aligned} \tag{4}$$

Thus, the problem is simplified to:

$$\min_y \sum_i^N \| \underbrace{\sum_z \left(x^{3D}(z) x_i^{2D}(\sigma_i) \right)^2 \left(x_i^{2D}(\sigma_i) - x^{3D}(z) \right)}_{f_i(y)} \|_2 \tag{5}$$

1.2. point-based representation

When X^{2D} and X^{3D} use point-based representations, the problem can be transformed into a nonlinear least squares problem from the reprojection error with weights as follows.

$$\min_{y,X^{3D}} \quad \sum_i^N \| w_i^{2D} \circ \left(\pi(F(x_i^{3D},y)) - x_i^{2D} \right) \|_2 \tag{6}$$

where $\pi(\cdot)$ denotes the linear transformation function obtained from the camera intrinsics, $F(\cdot)$ denotes the projection function obtained from the camera extrinsics, and $\| \cdot \|_2$ denotes the L2 norm.

*Corresponding author

First, we establish the constraint relationship between y and X^{3D} by projection and triangulation, and define $x^{2D} = F(x^{3D}, y), x^{3D} = G(x^{2D}, y)$ to simplify the problem. Then the above problem is transformed into:

$$\min_{y} \quad \sum_{i}^{N} \| \underbrace{w_i^{2D} \circ \left(\pi \left(F(G(x_i^{2D}, y), y)\right) - x_i^{2D}\right)}_{f_i(y)} \|_2 \tag{7}$$

1.3. KL

t(y) is the impulse function defined at the $y_g t$ point. The derivation of KL dispersion in train loss is as follows.

$$\begin{aligned}
\mathcal{L}_{cam} &= D_{KL}(z(y) \| p(y|X^{2D})) \\
&= \int z(y) \left(\log z(y) - \log p(y|X^{2D})\right) \mathrm{d}y \\
&= \int z(y) \log z(y) \mathrm{d}y - \int z(y) \log \frac{p(X^{2D}|y)}{E[p(X^{2D}|y)]} \mathrm{d}y \\
&= const - \int z(y) \log p(X^{2D}|y) \mathrm{d}y \\
&\quad + \log E\left[p(X^{2D}|y)\right] \int z(y) \mathrm{d}y \\
&= const - \log p(X^{2D}|y_{gt}) + \log E\left[p(X^{2D}|y)\right] \\
&= const - \frac{1}{2} \sum_{i}^{N} \|f_i(y_{gt})\|_2 + \log E\left[p(X^{2D}|y)\right]
\end{aligned} \tag{8}$$