Gaussian Label Distribution Learning for Spherical Image Object Detection

Supplementary Material

Hang Xu[1,2*] Xinyuan Liu[2] Qiang Zhao[2*†] Yike Ma[2] Chenggang Yan[1] Feng Dai[2†]
[1]Hangzhou Dianzi University, Hangzhou, China
[2]Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences,
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
{hxu, cgyan}@hdu.edu.cn, {liuxinyuan21s, zhaoqiang, ykma, fdai}@ict.ac.cn

Derivation details of coordinate transformation

We have modeled the rectangular tangent plane of spherical bounding box as a Gaussian distribution with some coordinate transformations in the main body of our paper, and thus now we provide more derivation details about these crucial transformations here.

For the sake of narration, we do not make a distinction between the *retangular tangent plane* and *spherical bounding box*. Before more discussion, we name the vector from the sphere center to the tangent plane center as *position vector*, i.e.

$$\begin{aligned}\mathbf{p}(\theta,\phi) &= [px, py, pz]^\top \\ &= [\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi)]^\top\end{aligned} \tag{1}$$

We start with the simplest case, i.e. rectangular tangent plane without rotation on the polar, denoted as $\mathbf{\Pi}(\theta, 0, w, h, 0)$. In this case, we can easily model the tangent plane as a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, where

$$\begin{aligned}\boldsymbol{\mu} &= \mathbf{p}(\theta, 0) \\ \mathbf{\Sigma} = \mathbf{\Lambda} &= \begin{bmatrix} \frac{w^2}{4} & 0 & 0 \\ 0 & \frac{h^2}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix}\end{aligned} \tag{2}$$

Here, the mean $\boldsymbol{\mu}$ referred to the center of Gaussian is set as *position vector*, and the covariance $\mathbf{\Sigma}$ referred to the anisotropic radius of Gaussian is set based on half of the orthometric edges of the tangent plane.

To get the Gaussian corresponding to the tangent plane with rotation in an arbitrary position, we apply two different coordinates to the basic Gaussian.

*This work was done when Hang Xu and Qiang Zhao were at ICT.
†Corresponding author.

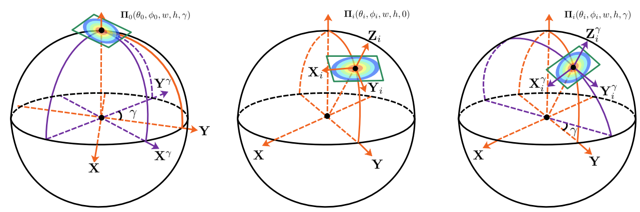

Figure 1. Schematic diagram of (c) modeling the rectangular tangent plane of a spherical bouding box as a Guassian distirbution based on (a) rotation transformation $\mathbf{R}$ and (b) movement transformation $\mathbf{T}$. **[Best viewed by zooming in.]**

(1) On the one hand, we construct a rotation transformation $\mathbf{R}$ based on rotation angle (γ) of the bounding box, as is shown in Fig. 1(a), making the Gaussian rotate around *position vector*, where

$$\begin{aligned}\mathbf{R} &= \left[\mathbf{X}^\gamma, \mathbf{Y}^\gamma, \mathbf{Z}^\gamma\right] \\ &= \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}\end{aligned} \tag{3}$$

(2) On the other hand, we construct another movement transformation $\mathbf{T}$ based on center position (θ_i, ϕ_i) of the bounding box, as is shown in Fig. 1(b), to make the Gaussian move to the specific position, where

$$\begin{aligned}\mathbf{T} &= \left[\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i\right] \\ &= \begin{bmatrix} \sin(\theta_i) & \cos(\phi_i)\cos(\theta_i) & \sin(\phi_i)\cos(\theta_i) \\ -\cos(\theta_i) & \cos(\phi_i)\sin(\theta_i) & \sin(\phi_i)\sin(\theta_i) \\ 0 & -\sin(\phi_i) & \cos(\phi_i) \end{bmatrix}\end{aligned} \tag{4}$$

In detail, we specify the *position vector* as z-axis in new coordinate system, i.e.

$$\hat{\mathbf{Z}} = \mathbf{p} \tag{5}$$

Moreover, since the *tangent vector* along the longitude θ is aligned with the height h of the rectangular tangent plane, we specify it as the y-axis, which is actually the partial derivative vector of *position vector* with respect to latitude ϕ, i.e.

$$\hat{\mathbf{Y}} = \frac{\partial \mathbf{p}}{\partial \phi} \tag{6}$$

Finally, we can get corresponding new x-axis aligned with the width w of the rectangular tangent plane, based on the cross product from $\hat{\mathbf{Z}}$ to $\hat{\mathbf{Y}}$, i.e.

$$\hat{\mathbf{X}} = \hat{\mathbf{Y}} \times \hat{\mathbf{Z}} \tag{7}$$

Based on the above two transformations $\mathbf{R}$ and $\mathbf{T}$, we can get the Gaussian corresponding the tangent plane with rotation in an arbitrary position, where

$$\begin{aligned} \boldsymbol{\mu}_i &= \mathbf{p}(\theta_i, \phi_i) \\ \boldsymbol{\Sigma}_i &= \mathbf{T}\boldsymbol{\Sigma}_0\mathbf{T}^\top = \mathbf{R}(\mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^\top)\mathbf{R}^\top \end{aligned} \tag{8}$$

In this case, actual coordinate system is based on $\mathbf{X}_i^\gamma$, $\mathbf{Y}_i^\gamma$, $\mathbf{Z}_i^\gamma$, as is shown in Fig. 1(c).