Unbiased IoU for Spherical Image Object Detection

Supplementary Material

Anonymous AAAI-22 Submission

Paper ID: 2266

In this appendix (supplemental material), we introduce the derivation of the formula for the area of spherical rectangle given its field of view α and β. We also describe the details of how to compute the radius used to generate the ground truth heatmap.

Appendix A: Area of Spherical Rectangle

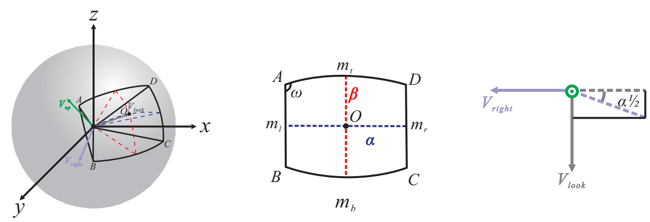
According to the definition of spherical polygons, we know that the four sides of each spherical rectangle are all great-circle arcs, which are the intersection of the surface with planes through the center of the sphere (Wikipedia contributors 2021). We also know that the four angles $\{\omega_i, i = 1, \cdots, 4\}$ of each spherical rectangle are equal due to symmetry. Therefore we can compute the area of a spherical rectangle b as $A(b) = 4\omega - 2\pi$ according to the formula for the area of spherical polygons, where ω is the angle of spherical rectangle and is defined as the angle between the planes that the neighboring sides of each spherical rectangle lie on. Thus, the key step is how to compute the value of angle ω.

The angle between two planes equals to the angle between the normals of these two planes. For each spherical rectangle, the four planes corresponding to its four sides can be obtained by rotating the planes that pass through the center point (θ, ϕ) of the spherical rectangle. Taking the left plane P_{Am_lB} as an example, as shown in Figure 1 (a) and (b), it can be obtained by rotating the plane $P_{m_tOm_b}$ around the axis V_{up} by $\frac{\alpha}{2}$, where α is the horizontal field of view of the spherical rectangle. The V_{up} is an axis of the coordinate system that we established based on axis $\vec{z} = [0, 0, 1]^T$ and the azimuthal and polar angle (θ, ϕ) of the center point of the spherical rectangle

$$\begin{cases} V_{look} = [\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi)]^T \\ V_{right} = [-\sin(\theta), \cos(\theta), 0]^T \\ V_{up} = [-\cos(\phi)\cos(\theta), -\cos(\phi)\sin(\theta), \sin(\phi)]^T \end{cases} \tag{1}$$

As a consequence, the normal of the left plane P_{Am_lB} can also be obtained by rotating the normal of plane $P_{m_tOm_b}$ around the axis V_{up} by $\frac{\alpha}{2}$. This is illustrated in Figure 1 (c), in which the dashed gray line and the dashed blue line denote

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.



(a) (b) (c)

Figure 1: (a) To compute the area of a spherical rectangle, we establish a local coordinate system $V_{look}V_{right}V_{up}$. (b) Then the planes corresponding to four sides of the spherical rectangle can be obtained by rotating the planes determined by the center point of the spherical rectangle. (c) The normals of the planes can be easily computed based on the established coordinate system, which can be used to compute the area.

the normal of the plane $P_{m_tOm_b}$ and P_{Am_lB} respectively. Thus, the normal of the left plane P_{Am_lB} can be derived as

$$N_l = \sin\frac{\alpha}{2}V_{look} - \cos\frac{\alpha}{2}V_{right}. \tag{2}$$

Similarly, we can also get the normal of the top plane as

$$N_t = \sin\frac{\beta}{2}V_{look} - \cos\frac{\beta}{2}V_{up}. \tag{3}$$

The angle ω is given by

$$\omega = \pi - \arccos(\langle N_l, N_t \rangle) = \arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2}) \tag{4}$$

where $\langle \cdot, \cdot \rangle$ is the dot product of two input vectors.

The normal of the right plane N_r and the normal of the bottom plane N_b is given in the similar way

$$N_r = \sin\frac{\alpha}{2}V_{look} + \cos\frac{\alpha}{2}V_{right}. \tag{5}$$

$$N_b = \sin\frac{\beta}{2}V_{look} + \cos\frac{\beta}{2}V_{up}. \tag{6}$$

Then we can calculate the four angles through Equation 4 and most interestingly find that they are all equal.

After obtaining the four angles, the area $A(b)$ of spherical rectangle b can be consequently computed as

$$A(b) = 4\arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2}) - 2\pi \tag{7}$$

Appendix B: The Calculation of Radius

When generating the ground truth heatmap, we assign non-zero values to the negative locations within a radius of the positive location as in Corner (Law and Deng 2018) and CenterNet (Zhou, Wang, and Krähenbühl 2019). Here we describe how to calculate the radius. Actually, there are three cases for the radius, which correspond to different relationships between ground truth and predicted bounding boxes as shown in Figure 2. We only derive for the first case in detail, and the other two cases can be derived in a similar way as follows.

Case a: Let α and β be the field of view of ground truth bounding box, and let γ be the radius. The IoU, whose threshold is t, between the predicted and the ground truth bounding box is

$$\frac{4\arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2})-2\pi}{4\arccos\left(-\sin(\frac{\alpha+2\gamma}{2})\sin(\frac{\beta+2\gamma}{2})\right)-2\pi}=t. \tag{8}$$

According to the product-to-sum identities of trigonometric functions,

$$\begin{aligned}&\sin(\frac{\alpha+2\gamma}{2})\sin(\frac{\beta+2\gamma}{2})\\ &=-\frac{1}{2}\left[\cos(\frac{\alpha+\beta+4\gamma}{2})-\cos(\frac{\alpha-\beta}{2})\right].\end{aligned} \tag{9}$$

Thus Equation 8 can be written as

$$\begin{aligned}&2\arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2})-\pi\\ =&t\left[2\arccos\left(\frac{1}{2}\left[\cos(\frac{\alpha+\beta+4\gamma}{2})-\cos(\frac{\alpha-\beta}{2})\right]\right)-\pi\right].\end{aligned} \tag{10}$$

The above equation can be reduced to

$$\begin{aligned}&\cos(\frac{\alpha+\beta+4\gamma}{2})\\ =&2\cos\left(\frac{\arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2})-\frac{\pi}{2}}{t}+\frac{\pi}{2}\right)+\cos(\frac{\alpha-\beta}{2}),\end{aligned} \tag{11}$$

then

$$\begin{aligned}\gamma=&\frac{1}{2}\arccos\left(-2\sin\left(\frac{\arcsin(\sin\frac{\alpha}{2}\sin\frac{\beta}{2})}{t}\right)+\cos(\frac{\alpha-\beta}{2})\right)\\ &-\frac{\alpha+\beta}{4}.\end{aligned} \tag{12}$$

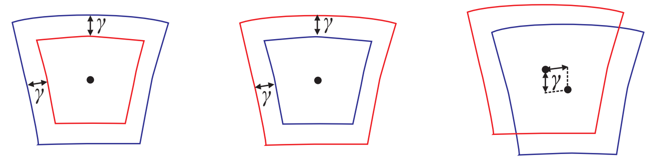
The derivation of **Case b** and **Case c** is the same as above, thus we only give the final radius calculating formulas.

Case b:

$$\begin{aligned}\gamma=-\frac{1}{2}\arccos\left(-2\sin\left(t\arcsin(\sin\frac{\alpha}{2}\sin\frac{\beta}{2})\right)+\right.\\ \left.\cos(\frac{\alpha-\beta}{2})\right)+\frac{\alpha+\beta}{4}.\end{aligned} \tag{13}$$

Case c:

$$\begin{aligned}\gamma=-\arccos\left(-2\sin\left(\frac{2t\left(\arccos(-\sin\frac{\alpha}{2}\sin\frac{\beta}{2})-2\pi\right)}{1+t}\right)+\right.\\ \left.\cos(\frac{\alpha-\beta}{2})\right)+\frac{\alpha+\beta}{2}.\end{aligned} \tag{14}$$



(a) (b) (c)

Figure 2: Different relationships between the ground truth (red) and the predicted (blue) bounding boxes: (a) the predicted bounding box contains the ground truth bounding box, (b) the ground truth bounding box contains the predicted bounding box, (c) these two bounding boxes intersect with each other.

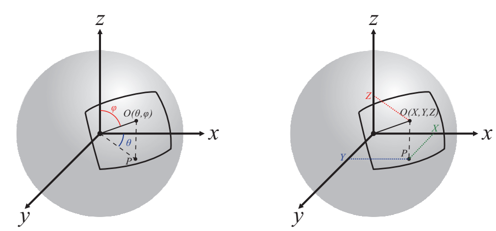
The final radius is the minimum of the above three cases, which is

$$\gamma=\min(\gamma_{\boldsymbol{a}},\gamma_{\boldsymbol{b}},\gamma_{\boldsymbol{c}}) \tag{15}$$

where $\gamma_{\boldsymbol{a}},\gamma_{\boldsymbol{b}},\gamma_{\boldsymbol{c}}$ represent the radius of the above three cases respectively.

Appendix C: Transformation from the azimuthal and polar angle to 3D unit vector

Transformation $\mathcal{T}(\cdot)$ that converts the azimuthal and polar angle to 3D unit vector in Cartesian coordinate system is required to use in the offset ground truth generation and prediction branch. Here we give a brief description of this conversation procedure.



(a) (b)

Figure 3: Transformation from the azimuthal and polar angle to 3D coordinate. (a) the azimuthal and polar angle coordinate representation, (b) the 3D coordinate representation.

The original center O of the spherical rectangle uses the azimuthal and polar angle (θ,φ) as the representation shown in Figure 3 (a). We need to transpose this representation to the 3D coordinate (X,Y,Z) as shown in Figure 3 (b) for more plausible loss computation. At first, the basic transformation formula can be written as

$$\begin{cases}X=\sin\varphi\cos\theta\\ Y=\sin\varphi\sin\theta\\ Z=\cos\varphi\end{cases} \tag{16}$$

According to Equation 16, the 3D coordinate of the center point O can be calculated, and the vector form is $\vec{O}=[X,Y,Z]$. Then the normalization operation needs to be performed in order to produce the unit coordinate by

$$\begin{cases} X' = \frac{X}{\|\vec{O}\|_2} \\ Y' = \frac{Y}{\|\vec{O}\|_2} \\ Z' = \frac{Z}{\|\vec{O}\|_2} \end{cases} \tag{17}$$

where $\|\vec{O}\|_2$ is given by

$$\|\vec{O}\|_2 = \sqrt{X^2 + Y^2 + Z^2} \tag{18}$$

Through the above derivation, we can get the 3D unit vector of the center point O of the spherical rectangle: $\vec{O} = [X', Y', Z']$. Please note that this derivation process can be used for the 2D angle coordinate to 3D unit vector transformation of any point on the sphere. With the 3D unit vector, the following dot product of two vectors and the final offset loss can be easily calculated.

References

Law, H.; and Deng, J. 2018. CornerNet: Detecting Objects as Paired Keypoints. In *European Conference on Computer Vision*, 765–781.

Wikipedia contributors. 2021. Spherical trigonometry. https://en.wikipedia.org/w/index.php?title=Spherical_trigonometry&oldid=1016967508.

Zhou, X.; Wang, D.; and Krähenbühl, P. 2019. Objects as points. *arXiv:1904.07850*.