Network-agnostic distortion-robust projections
for wide-angle image understanding - Supplementary Material

Akshaya Athwale[1], Ola Ahmad[2], Jean-François Lalonde[1]

[1]Université Laval   [2]Thales CortAIx Labs Canada

1. Introduction

In this supplementary material, we provide additional details to support the findings and methods presented in the main paper. Specifically, we include:

- Sec. 2 explains training and evaluation metric used in experiments.
- Sec. 3.1 explains equations following up on sec. 4.3 of the main text to determine which face of the cubemap the coordinates (u, v) belongs to.
- Sec. 3.2 further explains HEALPix grid as explained in sec. 4.4 of the main text, including its north polar cap and north equatorial belt.
- Sec. 4 explains the inverse warping corresponding to each forward warping described in the sec. 4 of main text.

2. Training loss, evaluation metrics and training images

We use the cross entropy and dice loss functions for training the networks

$$\ell_{\text{CE}}(y, \hat{p}) = -\sum_{n=1}^{N} \log(\hat{p} \cdot y) , \quad (1)$$

$$\ell_{\text{dice}}(y, \hat{p}) = 1 - \frac{(2y\hat{p} + 1)}{(y + \hat{p} + 1)} , \quad (2)$$

where $\hat{p}$ is the predicted probability and y is class labels. The total loss ℓ is defined as:

$$\ell = 0.4\ell_{\text{CE}} + 0.6\ell_{\text{dice}} . \quad (3)$$

These hyperparameters were adopted according to the specifications in [1]. Performance is evaluated using the mean Intersection-over-Union (mIoU) metric.

3. Projection equations

3.1. Cubemap

Based on the grid coordinates of cubemap (u, v), we determine which face it belongs, using the AND operator $\wedge$, and

Very low $\xi \in [0.0, 0.05]$ Low $\xi \in [0.2, 0.35]$ Medium $\xi \in [0.5, 0.7]$ High $\xi \in [0.85, 1.0]$
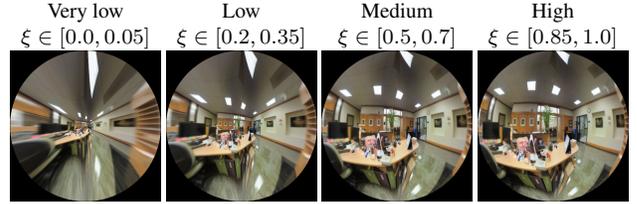
Figure 1. Visualization of wide-angle images at 4 different distortion levels, from left to right: very low, low, medium, and high.

compute the corresponding world coordinates (X, Y, Z).
Up : $\wedge(\wedge(u >= 1, u < 2), v < 1)$,

$$X = (u - 1.5) \times 2; Y = 1; Z = (v - 0.5) \times -2 , \quad (4)$$

Left : $\wedge(\wedge(u < 1, v >= 1), v < 2)$,

$$X = -1; Y = (v - 1.5) \times -2; Z = (u - 0.5) \times -2 , \quad (5)$$

Forward : $\wedge(\wedge(\wedge(u >= 1, u < 2), v >= 1), v < 2)$,

$$X = (u - 1.5) \times 2; Y = (v - 1.5) \times -2; Z = -1 , \quad (6)$$

Right : $\wedge(\wedge(u >= 2, v >= 1), v < 2)$,

$$X = 1; Y = (v - 1.5) \times -2; Z = (u - 2.5) \times 2 , \quad (7)$$

Down : $\wedge(\wedge(\wedge(u >= 1, u < 2)v >= 2), v < 3)$,

$$X = (u - 1.5) \times 2; Y = -1; Z = (v - 2.5) \times 2 . \quad (8)$$

3.2. HEALPix

Locations on the healpix sphere are defined by $(cos(\theta), f_\varphi)$ where $\theta \in [0, \pi]$ is the colatitude in radians measured from the North Pole and $f_\varphi \in [0, 2\pi]$ is the longitude in radians measured Eastward.

Pixel centers on the northern hemisphere are given by the following equations:

North polar cap For $p_h = (p + 1)/2$, the ring index $1 \leq i < N_{\text{side}}$, and the pixel-in-ring index $1 \leq j \leq 4i$,

$$i = I\left(\sqrt{p_h} - \sqrt{I(p_h)}\right) + 1 \tag{9}$$

$$j = p + 1 - 2i(i - 1) \tag{10}$$

$$z = 1 - \frac{i^2}{3N_{\text{side}}^2} \tag{11}$$

$$\varphi = \frac{\pi}{2i}\left(j - \frac{s}{2}\right), \quad \text{and} \quad s = 1 \tag{12}$$

North equatorial belt For $p' = p - 2N_{\text{side}}(N_{\text{side}} - 1)$, $N_{\text{side}} \leq i \leq 2N_{\text{side}}$, and $1 \leq j \leq 4N_{\text{side}}$,

$$i = I(p'/4N_{\text{side}}) + N_{\text{side}} \tag{13}$$

$$j = (p' \mod 4N_{\text{side}}) + 1 \tag{14}$$

$$z = \frac{4}{3} - \frac{2i}{3N_{\text{side}}} \tag{15}$$

$$\varphi = \frac{\pi}{2N_{\text{side}}}\left(j - \frac{s}{2}\right), s = (i - N_{\text{side}} + 1) \mod 2 \tag{16}$$

where the auxiliary index s describes the phase shifts along the rings, and $I(x)$ is the largest integer number smaller than x. Pixel center positions on the Southern hemisphere are obtained by the mirror symmetry of the grid with respect to the Equator ($z = 0$).

4. Inverse projections

In this section, we provide a detailed description of the inverse warping methods, corresponding to the warping methods explained in sec. 4 of the main text.

4.1. Preliminaries

For inverse warping, we aim to warp the target projection from coordinates (u, v) back to the wide-angle image with coordinates (x, y). We begin with a grid of coordinates in the wide-angle image space (x, y) and compute the corresponding target warp image coordinates (u, v) through inverse projection. This allows us to retrieve pixel values from the target image using bilinear sampling, ensuring smooth and accurate warping. The warping process follows these steps:

1. **Compute incident and azimuth angles** (θ, φ). For each wide-angle image coordinate (x, y), the incident angle can be obtained by first computing $r_w = f_{\text{lens}}(\theta)$, the radial distance of the pixel in the wide-angle image using,

$$r_w = \sqrt{x^2 + y^2}, \tag{17}$$

where $f_{\text{lens}}(\theta)$ depends on the lens distortion function used. The azimuth angle $\varphi = f_\varphi(x, y)$ is also determined as,

$$\varphi = \tan^{-1}\left(\frac{y}{x}\right) \tag{18}$$

The functions $(f_r, f_\varphi, f_{\text{proj}})$ might use the above equations or are projection-specific and will be defined in the following sections.

2. **Compute target coordinates** (u, v). Using the computed incident angle θ, we determine the radius in the target image r via the projection function, $r = f_{\text{proj}}(\theta)$. This radius, along with the azimuth angle φ, are then used to compute the corresponding wide-angle image coordinates (x, y) as

$$u = r\cos\varphi \text{ and } v = r\sin\varphi. \tag{19}$$

Finally, bilinear sampling is applied to map the distorted image onto the target grid.

4.2. Inverse polar projection

We start with a wide-angle image coordinate grid (x, y) and calculate θ using the lens distortion projection $\theta = f_{\text{lens}}^{-1}(r_w)$, where r_w is calculated using eq. (17). We then compute f_φ using eq. (18) and determine the target image coordinates (u, v) using eq. (19).

4.3. Inverse cubemap projection

We start with a wide-angle image coordinate grid (x, y) and calculate θ using the lens distortion projection $\theta = f_{\text{lens}}^{-1}(r_w)$, where r_w where is calculated using eq. (17). Using θ and depending on the wide-angle lens projection, we calculate the (X, Y, Z) coordinates.

We do not calculate the azimuth f_φ; instead, we map the world coordinates (X, Y, Z) onto the cubemap coordinates (u, v) using perspective projection. To determine which world coordinate lies on which face of the cubemap, we use the following equations.

Forward face : $\wedge(\wedge(Z <= 0, Z <= - \mid X \mid), Z <= - \mid Y \mid)$.

$$u = 1.5 - 0.5 * X/Z$$
$$v = 1.5 + 0.5 * Y/Z$$

Down face : $\wedge(\wedge(Y <= 0, Y <= - \mid X \mid), Y <= - \mid Z \mid)$.

$$u = 1.5 - 0.5 * X/Y$$
$$v = 2.5 - 0.5 * Z/Y$$

Up face : $\wedge(\wedge(Y >= 0, Y >= \mid X \mid), Y >= \mid Z \mid)$.

$$u = 1.5 + 0.5 * X/Y$$
$$v = 0.5 - 0.5 * Z/Y$$

Left face : $\wedge(\wedge(X <= 0, X <= - \mid Y \mid), X <= - \mid Z \mid)$.

$$u = 0.5 + 0.5 * Z/X$$
$$v = 1.5 + 0.5 * Y/X$$

Right face : $\wedge(\wedge(X >= 0, X >= \mid Y \mid), X >= \mid Z \mid)$.

$$u = 2.5 + 0.5 * Z/X$$
$$v = 1.5 - 0.5 * Y/X$$

4.4. Inverse elliptical grid mapping (EGM) projection

In the inverse EGM projection, before computing the incident θ, we start with the wide-angle coordinates (x, y) and calculate the intermediate coordinates,

$$x' = \frac{1}{2}\sqrt{2 + \alpha + 2\sqrt{2}x} - \frac{1}{2}\sqrt{2 + \alpha - 2\sqrt{2}x}, \quad (20)$$

$$y' = \frac{1}{2}\sqrt{2 - \alpha + 2\sqrt{2}y} - \frac{1}{2}\sqrt{2 - \alpha - 2\sqrt{2}y}, \quad (21)$$

as shown in fig. 3 of main text, where $\alpha = x^2 - y^2$.

Now we, calculate incident angle θ using $r_w = f_{\text{lens}}(\theta)$. Unlike other projection, we don't compute r_w using eq. (17), we use L_∞ norm radius where $r_w = \max(|x'|, |y'|)$.

We calculate target warp coordinate (u, v) using $r = f_{\text{proj}}(\theta)$, we define f_{proj} as explained in sec. 4.2 of main paper. The azimuth angle φ is not required and the wide-angle coordinates are not computed using eq. (19). Rather, we compute target image coordinates (u, v) (see fig. 3 in main text) by considering the set of all points at an L_∞ distance of r from the origin. This set forms a square with eight key points, defined as:

$$u, v \in \{(\pm r_w, 0), (0, \pm r), (\pm r, \pm r)\} . \quad (22)$$

The eight key points consist of the four corners and four edge midpoints of a square. These points form concentric squares, ensuring a structured and uniform grid completion.

4.5. Inverse HEALPix projection

We first calculate the incident angle $\theta = f_{\text{lens}}^{-1}(r_w)$, where r_W is calculate using eq. (17) and f_φ is calcualte usign eq. (18).

We calculate the healpix grid coordinate $(u, v) \equiv (\cos\theta, \phi)$ [2].

References

[1] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-Unet: Unet-like pure transformer for medical image segmentation. In *Eur. Conf. Comput. Vis.*, 2022. 1

[2] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astro. J.*, 622(2): 759–771, 2005. 3