Project 3 FAQ

Assumption

We assume there's no big change in the tilting angle of camera pose in this project. Under this assumption, you only need to apply feathering (feathering width (or blending width) is in pixels) to the left end and right end of the images (no need to do it for upper and lower end). The feathering procedure is applied first to each image before accumulating it to the accumulated image.

Image Accumulation and Normalization

Accumulating procedure is simply an online way to compute the weighted average of all the **RGBA** images that are to be stitched (the **A** channel is computed by feathering). See below (I denotes RGB image, with α being the alpha channel). Basically, you would need to keep track of the numerator and denominator whenever a new image is added. The numerator could be stored in the **RGB** channel of the accumulated image, with the denominator in the **A** channel. The final division step is called normalization in the skeleton code. (Reference slide 54 of the lecture.)

$$I = \frac{\sum_{i=1}^{N} \alpha_i I_i}{\sum_{i=1}^{N} \alpha_i}$$

Inverse Warping

Good resource to refer to
https://blogs.mathworks.com/steve/2006/05/05/spatial-transformations-inverse-mapping/

Representing a Light Ray in Camera Coordinate Frame

In camera coordinate frame, there's a one-to-one correspondence between the imaged 3D point $(X, Y, Z),\ Z > 0$, and the point on the normalized image plane $Z = 1$; specifically,

$$(X, Y, Z) \;<\!\!-\!\!-\!\!-\!\!>\; (X/Z,\; Y/Z,\; 1)$$

The physical intuition is that the light ray which $(X, Y, Z),\; Z > 0$ lies on always intersects with the image plane once.

Alternatively, this light ray also intersects with the unit hemispherical surface once, which means you could establish a one-to-one correspondence between a 3D point $(X, Y, Z),\; Z > 0$ and a point on the unit hemispherical surface. To represent a point on this unit hemispherical surface, you need two angles $(\theta,\; \phi)$. The two angles are defined in the following way.

- θ: angle between the plane defined by **(Y axis, Z axis)**, and the plane defined by **(Y axis, the light ray)**.
- ϕ: angle between **the light ray** and it's **orthogonal projection on the X-Z plane.**

Under this definition, you would have the following one-to-one correspondence.

$$(X, Y, Z) \;<\!\!-\!\!-\!\!-\!\!>\; (\cos\phi\sin\theta,\;\; \sin\phi,\;\; \cos\phi\cos\theta)$$

This type of representation is especially convenient when you try to stitch images taken by a rotating camera, as the rotation turns into translation in the $(\theta,\; \phi)$ coordinates.

Spherical Image

A digital image is a discretization of the projected points on the image plane, i.e., $(X/Z,\; Y/Z,\; 1)$. This discretization is done via the intrinsic matrix. Note that since the unit of focal length f in the intrinsic matrix is pixels, the discretization has a resolution such that each pixel roughly covers $1/f$ angular range in radians. To get the spherical image expected in this project, you would need to apply the same discretization procedure to the (θ, ϕ) coordinates, with the same resolution $1/f$.