Coordinate Systems in 3D-PTV Algorithms

Yosef Meller

July 12, 2016

The 3D-PTV method of finding 3D positions of particles in flow has an easy to understand general scheme: use 2D images of particles taken from several distinct points of view for calculating a 3D position of the observed particles. The imlementation of this idea, however, is somewhat more involved. The basic idea already implies the existence of at least 2 coordinate systems - the 2D view coordinates and the global 3D coordinates. Practical considerations expand this list further, and this document attempts to list and summarize the purpose and properties of all coordinate systems involved in the process. A more thorough but possibly less accessible guide may be found in ref. [1].

1 Spatial Coordinates

There are two types of spatial (3D) coordinate systems involved in 3D-PTV: the Global Coordinates, and each camera's Local Frame.

The *Global Coordinates* are the base coordinate system. Everything in the PTV system lies within it. Although this system is initially arbitrary, it is commonly expressed in millimetres for small to medium experiments and is determined by the arbitrarily defined positions of points on a calibration target. Typically one well-identified point on the target will be designated as the origin, and two other points (in an orthogonal point spread) or even just one (non-orthogonal) will determine the direction of axes. It is important, when designating the control points, to note that the Z direction must be consistent with a right-handed system, and that all other calibration points should be consistent with the determined axes.

The *Local Frame* of each camera is received from the global coordinates by rotation and translation. Their origin is the camera's *primary point* (The imaginary focal

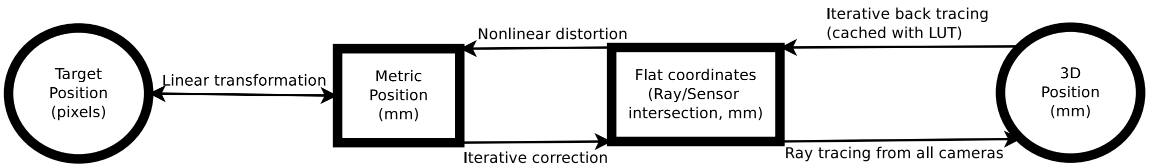

Figure 1: Image coordinate systems

point of the Tchen camera model, which resembles a pinhole camera). The Z axis points opposite the lens direction (i.e. the sensor and observed volume are always in the local negative Z). The local X and Y directions are aligned with the image, such that if the Z axis is horizontal in the global frame and the camera is held upright, the Y direction points up. The X axis is right-handed and therefore would in this setting point to the right of an imaginary photographer holding the camera.

In OpenPTV The image plane position is described in terms of the local coordinates, and 3D operations such as ray tracing and finding image coordinated rely on this description together with the transformations that describe the frame. In some other PTV experiments, the Tchen model is discarded in favor of a fitted function transforming image coordinates to 3D rays directly in global coordinates. Then the local frame is not needed.

2 Image Coordinates

The most relevant image coordinate systems are the native Pixel Coordinates, the Metric Coordinates, and the Flat Coordinates, which account for camera distortions. Their relarelationship is described in this section, and summarized in Figure .

The *Pixel Coordinates* are the most obvious - they are the row and column in the image data matrix. The origin is at the top-left of the image, the y axis points down, the x axis points right, and the units are pixels. The *Metric Coordinates* are a simple linear transformation of this system into one where the origin is at the *image*'s center point, y points up, x still points right, and units are in millimetres. The unit conversion factor is based on the pixel size in the sensor, so that the rightmost pixel x-coordinate in the Metric system is half the sensor width in millimetrs. The sensor width and pixel size may be found in camera datasheets.

For example: For a sensor of 1280×1024 pixels (on the x and y direction respectively), each 0.014 mm to a side, Table 1 shows the pixel and metric coordinates of the image corners.

Corner	Pixel Coordinates (x, y)	Metric coordinates (x, y)
top left	0,0	-8.96,7.168
top right	1280,0	8.96,7.168
bottom left	0,1024	-8.96,-7.168
bottom right	1280,1024	8.96,-7.168

Table 1: Example of metric coordinates and the corresponding pixel coordinates.

Flat coordinates are a special case of the metric coordinates. They arise from the fact that lens distortion and sensor shift lead to a recorded image which is somewhat different than what would be seen by an ideal pinhole camera as assumed by the simpler model. The metric image coordinates denote coordinates of objects as seen by the camera. The flat coordinates denote the position where an ideal camera would see the same objects.

The flat coordinates are what you receive from using the multimedia code to trace back a ray from a known 3D position to its intersection with the sensor plane. To get to the Metric system, you must first add the sensor shift to the coordinates, then calculate the distorted coordinates using the usual distortion formulas.

The reverse operation is to find a point on the image plane that intersects the ray that goes from the camera primary point to the object represented by a given Metric target coordinates. This operation appears mainly in calculating average 3D position from ray intersections. To do this one must first *undistort* (or *correct*) the Metric coordinates. This is an inverse problem to that of distortion and is solved iteratively. Then the sensor shift is subtracted from the result to yield the Flat coordinates.

In the old 3DPTV code, with some exceptions, Pixel coordinates are held in the `pix` arrays; Metrc coordinates are in the `crd` arrays; and Flat coordinates are in the `geo` arrays.

References

[1] T. Dracos. *Three-Dimensional Velocity and Vorticity Measuring and Image Analysis Techniques: Lecture Notes from the Short Course held in Zürich, Switzerland, 3–6 September 1996.* ERCOFTAC Series. Springer Netherlands, 2013.