From Camera Parameters to World Coordinates

Image Processing Basics: Camera Models

Applications of image processing in robotics or 3D metrology often call for the determination of coordinates for points in real space, the so-called world-coordinates, from the data in the image file. The corresponding calculations are based on a quantitative geometric camera model. In contrast to radiometric models, the signal path is not relevant in this context. A geometric camera model describes the image formation from points in the workspace to pixel coordinates in the data file.

Central Projection

An image of a scene in three-dimensional space taken with a standard lens will be formed in central projection, resulting in warped geometric objects due to perspective distortion. Figure 1 shows an example. All the rectangular stones have the same edge-length in reality, of course, but appear more and more shrunken with increasing distance. In the upper half of figure 2, the situation is described within the so-called "pinhole camera model", which holds in good approximation for imaging with a standard lens with fixed focal length. The objects in the workspace are drawn on the right hand side of the lens. The detector chip is mounted in the image plane of the camera, shown on the left hand side of the lens. Two objects with identical dimensions, placed at different distances from the lens, will appear as images of different sizes due to central projection. A sharp image will appear for a single defined working distance only. In practical applications, however, the depth of field usually can be tuned to be sufficiently large by proper choice of the focal length and the f-number. The quantitative relations can be derived from the lower part of figure 2. A point with world-coordinates X_W und Z_W in the working space will be imaged to a point with the sensor-coordinate x_s in the vertical distance b from the centre of the projection. The parameter b is an internal parameter of the camera model, the image distance, sometimes (but not strictly correctly) denoted as the focal length. As can easily be seen, $X_W/Z_W=x_s/b$. This relation also holds, when the image is blurred. In this situation, the image of a point just grows to a small disk, but the position of this blob remains the same as in the corresponding sharp image within good approximation. In general, a point of interest in the scene will not necessarily be located in the X_W-Z_W-plane shown in figure 2 but may have a world-coordinate Y_W not equal to zero perpendicular to the drawing plane. The corresponding point in the image will have a sensor coordinate y_s not equal to zero. The central projection of a point with the world coordinates X_W,Y_W and Z_W to an image point with the sensor coordinates x_s and y_s is thus described by the following two equations:

Fig. 1: Warping by central projection

$$x_s = -b\,(X_w/Z_w)$$
$$y_s = -b\,(Y_w/Z_w) \qquad (1)$$

All quantities in these equations have linear units such as millimeter or micrometer. The sensor coordinates x_s und y_s, also measured in millimeters or micrometers, are not yet related with the pixel coordinates of the image in the data file, but are coordinates in the real world, although not in the workspace, but in the image plane of the camera. The pixel-coordinates of the data file have to be linked with real dimensions in the sensor plane by the well-known, precise array cell structure of the detector chip.

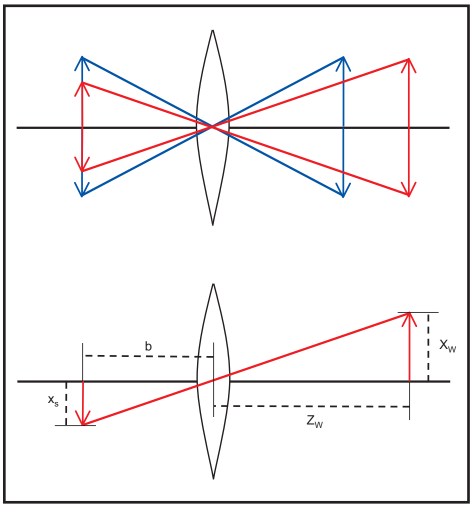

Fig. 2: Central projection in the pinhole-camera model

External and Internal Parameters

Geometric camera models use two different world-coordinate systems. The world-coordinate system of the camera already appeared in figure 2. In this system, a point in the workspace has the world-coordinates X_W, Y_W and Z_W, with the origin in the projection centre of the lens. The Z_W-axis is directed along the optical axis of the lens. This coordinate system will rotate and move with the camera when the orientation and the position of the camera (the so-called pose) are changed. Usually, however, it will be more convenient to use a second world-coordinate system in the workspace, which may be related to a fixed plane such as a conveyor belt or a tabletop. We denote these world-coordi-

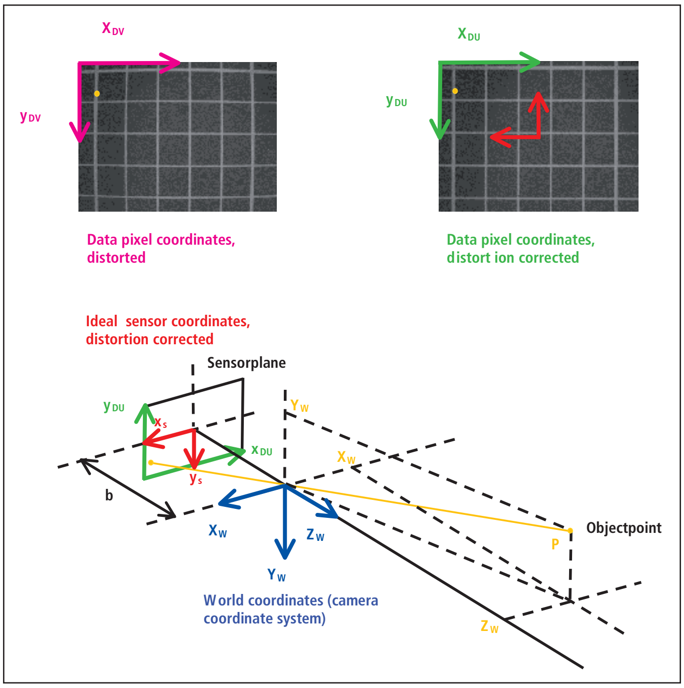

Fig. 3: Coordinate systems for the transformation of world-coordinates to pixel-coordinates in the image data file and the influence of optical distortion

nates as (X, Y, Z). This coordinate system may be transformed into the camera coordinate system (X_W, Y_W, Z_W) by two simple operations. The first operation is a translation of the origin of the system of the workspace into the origin of the camera coordinate system, characterized by the three components of the corresponding translation vector. The second operation is a rotation of the coordinate system such that the corresponding axes of the two systems point into the same directions. This rotation may be described by a rotation matrix with three independent parameters. The six degrees of freedom of the orientation and position of the camera in space are thus determined by six parameters, the so-called external (or exterior or extrinsic) camera parameters. The camera model now contains these six external parameters and the image distance as an internal (or interior or intrinsic) parameter.

The final step of the camera model is the connection with the pixel coordinates in the data file. The origin of the pixel coordinate system usually is placed in the upper left corner of the image which will be seen when viewing into the workspace from the position of the camera. The lens will image this point to the lower left corner of the detector chip, now viewed from the workspace back to the camera. Have a look at figure 3 to get a better insight into these somewhat complex geometric relations. The yellow object point has the world coordinates (X_W,Y_W, Z_W) in the blue world-coordinate system of the camera. The coordinate system (X, Y, Z) in the workspace is not drawn to avoid confusion. The object point is imaged to a point in the sensor plane, which will appear in the upper left corner of the image in the data file. To emphasize this point, the upper part of figure 3 contains a view of the computer screen where the data file is displayed. In the lower part, the green coordinate system of the data file is embedded in the sensor plane. In addition to the data file coordinate system (x_D, y_D), the camera model defines the red sensor coordinate system (x_s, y_s). The origin of this system is located at the point where the optical axis hits the detector plane, the so-called principal point. The x- and y-axes of this system are directed parallel to the X_W- and Y_W-axes of the camera world coordinate system. The coordinates x_s and y_s are measured in linear units such as millimeter or micrometer. The first step to get sensor coordinates from pixel coordinates is to shift the origin of the coordinate system of the data file to the centre point of the sensor. These centred coordinates are then multiplied by the edge lengths S_x and S_y along the x- und y-direction. The result will be sensor coordinates scaled with real-world units. The scaling factors S_x and S_y usually are taken from the data sheet of the camera. Cameras with a digital interface will transfer the data from the sensor into the image data file as they are, whereas with analog cameras the ratio between sampling frequency of the frame grabber and the pixel clock has to be taken into account. Pixel coordinates can be calculated from sensor coordinates according to the following equations:

$$x_D = - x_s/S_x + H_x$$
$$y_D = - y_s/S_y + H_y \qquad (2)$$

H_x and H_y are the pixel coordinates of the principal point. Cameras designed for industrial image processing usually will have fairly well adjusted sensors, and the principal point will simply be the centre of the chip in good approximation. The number of pixels on the x- and y-axis, respectively, will thus determine the values for H_x and H_y. These parameters, however, may also be treated as further degrees of freedom in the camera model.

Distortion Correction

For real-world applications, an additional modification of the camera model is mandatory, namely, the distortion correction. The well-known workhorses of industrial imaging are simple C-mount- or CS-mount lenses, which suffer from considerable optical distortions. An example is shown in the upper left corner of figure 3. The main part of the distortion has rotational symmetry and is a function of the radial distance from the optical axis only. A single parameter κ is sufficient to characterize this relation. One possible modeling uses the following equations:

$$x_{DU}^* = x_{DV}^*/(1+\kappa\, r_V^2)$$
$$y_{DU}^* = y_{DV}^*/(1+\kappa\, r_V^2) \qquad (3)$$

where $r_V^2 = x_{DV}^{*2} + y_{DV}^{*2}$. The parameter κ describes the transformation from undistorted to distorted pixel coordinates in the data file, but with reference to the principal point, marked by the asterisk. In the upper right part of figure 3 the result of the distortion correction is shown. The red sensor coordinates in figure 3 are thus the ideal, undistorted sensor coordinates, which would appear in the sensor plane by imaging with an ideal, distortion-free lens. The camera model now contains further five internal parameters.

Final Remarks

The camera model described in this article, with six external and six internal parameters, is based on the work of Lenz [1] and Tsai [2], published already in 1987. The numerical determination of the values of these parameters by so-called camera calibration, based on images of calibration targets with reference points at precisely known world coordinates, is mathematically quite complex. Modeling the imaging process in terms of linear algebra with matrix operations will greatly simplify any approach to these methods. Fortunately, there is a lot of literature on this topic with good explanations and documentation [3, 4] for those readers who would like to tackle these interesting problems.

References

[1] Lenz, R., Informatik-Fachberichte 149: Mustererkennung 1987, p. 212-216, Springer 1987
[2] Tsai, R. Y., IEEE Journal of Robotics and Automation, Vol. RA-3 (4) 1987, p. 323-344
[3] Steger, C., Ulrich, M., Wiedemann, C., Machine Vision Algorithms and Applications, Wiley-VCH, Weinheim 2008, p. 180 ff.
[4] Azad, P., Gockel, T., Dillmann, R., Computer Vision, Elektor, Aachen 2007, p. 73 ff.

▶ Author
Prof. Dr. Christoph Heckenkamp
Darmstadt University of Applied Sciences
Department of Optical Technology and Machine Vision
heckenkamp@h-da.de
www.fbmn.h-da.de