World Coordinates under Control

Image Processing Basics: Camera Calibration

Applications of image processing in robotics or 3D metrology often call for the determination of real-world coordinates from the 2D data in the image file. The relationship between image coordinates and world coordinates can be described by a quantitative model of the image formation process. Geometrical camera calibration provides the values of these parameters by analysis of the images of calibration targets with well-defined reference points. The article describes the principles of this method.

Central Projection

A robot for a pick-and-place-application, supported by a vision system, needs the coordinates of the object to be picked and of the target area where the object is to be delivered. The control unit of the robot will be fed with coordinates in real space, so-called world coordinates, with reference to a fixed point in space. Let us first have a look at the somewhat simplified situation of figure 1, where the workspace is a plane and all points of interest are within this plane. An image of this scene taken with a standard lens will be formed in central projection, resulting in warped geometric objects due to perspective distortion. Scaling in such an image is difficult, since a single constant factor is not sufficient to transform distances within the image plane to corresponding distances in a plane in workspace. This becomes evident when the optical axis is tilted as in figure 2, where all the rectangular stones in reality have the same edge-length, of course, but appear more and more shrunken with increasing distance.

Coordinate Systems

The image formation process for a point from the workspace to the image plane of the camera can be described by a matrix transform, the so-called camera model. The image data in the memory of the host are pixel coordinates x_D and y_D, without any dimensional units. The known edge-length of a detector pixel, however, provides the connection with the sensor coordinates x_S und y_S of the image in the senor plane, measured in micrometers or other real-world units. The central projection of an object point to the sensor plane is described best in the camera coordinate system. In this system, a point has the world coordinates X_W, Y_W und Z_W, with the origin in the projection centre of the lens. In our pick-and-place-scenario, however, it will be more convenient to use world coordinates just in the working plane of the robot. The X- and Y-axes of this system are embedded in the plane, the Z-axis being perpendicular to the plane. The Z-coordinate of all points in this plane will thus be zero and will be known for all points in the plane. Figure 3 shows the different coordinate systems. The complete transformation from the world coordinates (X, Y, Z) of an object point to the pixel-coordinates (x_D, y_D) of the corresponding image point in the image data file may be mathematically described by a single matrix operation. This mathematical camera model contains a number of parameters, such as the six degrees of freedom for the orientation of the camera in space, the focal length of the lens and the distortion parameter of the lens. Once the numerical values of the model parameters are known, the matrix operation can be inverted, and world coordinates may be calculated from pixel coordinates. There is an important restriction, however: pixel-coordinates generally only give access to the ratios (X_w/Z_w) and (Y_w/Z_w), but not to the absolute values of X_w, Y_w und Z_w themselves. Since two points in space which are on the

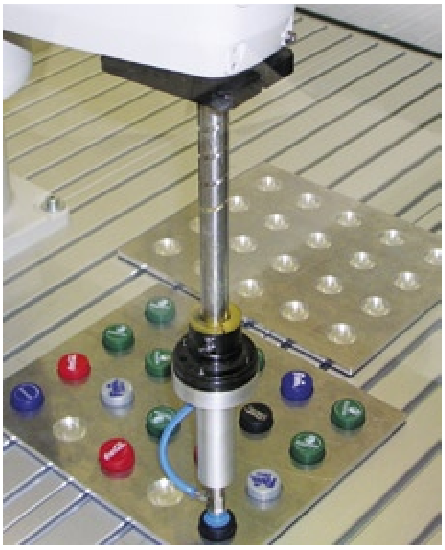

Fig. 1: Pick and place in a plane

same line of sight of the camera will inevitably be imaged to the same point on the detector, only the direction can be inferred from the pixel coordinates, the distance is lost in perspective. Whenever one of the three world coordinates is known, however, the other two may be calculated from the ratios.

Calibration

The transformation of the world coordinate system from workspace to the pixel coordinate system in the image data file is parameterized by 12 or more parameters, depending on the sophistication of the model. In general, these parameters, like the focal length or the distortion parameter, e.g., may be determined by separate and independent measurements. Common practice, however, is the calibration of the system by means of a calibration target. For this purpose, the camera in question takes images from an object with prominent, clearly visible features with well-known world coordinates in the coordinate system of the workspace. The positions of the corresponding points in the image are ex-tracted by the methods of image processing. After this operation, the world coordinates of the reference points in real space as well as the corresponding pixel coordinates are known. The link between both data sets is provided by the camera model. Since both sides of the matrix operation are known, the camera parameters may be treated as the unknown variables in the equations. With a sufficient number of reference points, there will be a system of equations which may be solved for the camera parameters. In our laboratory, we use a flat calibration target tailored to the pick-and-place-scenario of figure 1, which can be placed directly in the working plane of the robot. A suitable target is shown in figure 4. It contains circular objects with excellent contrast to the background, forming a regular mesh pattern with precise cell dimensions, well-known in absolute units. The origin and the orientation of the coordinate axes can easily be determined automatically. The huge circle in the centre defines the origin, and the small circle and the ring structure define the orientation of the X- and Y-axis, respectively. The Z-axis is directed perpendicular to the plane. The position of the circles in the image file is calculated as the centre of mass of the segmented objects after binarization and labeling of the image. The data for the positions obtained by this method are stable and have subpixel-precision. The requirements for homogeneous lighting are easily met with standard equipment.

Camera Model

The regular structure of the test pattern supports the automatic process of establishing the relationship be-tween the circles on the target with their well-known world coordinates in the system of the workspace and the corresponding centers of mass calculated from the image data. In figure 4, e. g., an array of six by six circles may safely

Fig. 2: Warping by central projection

be used to correlate world coordinates and pixel coordinates, resulting in correspondences for 36 well-known points in real space. The system of equations for the 12 camera parameters is thus significantly overdetermined and has to be solved with adequate numerical methods. The values for some parameters, the pixel dimensions, e. g., can be found in a data sheet with sufficient precision. For the remaining parameters, an optimized set of values will be determined by means of a least-square-fit. Suitable mathematical methods may be cumbersome, but are well documented and numerically stable. Nevertheless, the result of such calculations will be a set of parameters, which gives numerical access to the image formation according to the camera model. Since one of the three world coordinates is constant and known in the pick-and-place-plane of the robot, namely the Z-coordinate Z=0 for all points in the plane, the remaining two coordinates X und Y may now be calculated for every point in the plane from the pixel coordinates of the corresponding point in the image data file by means of the camera model.

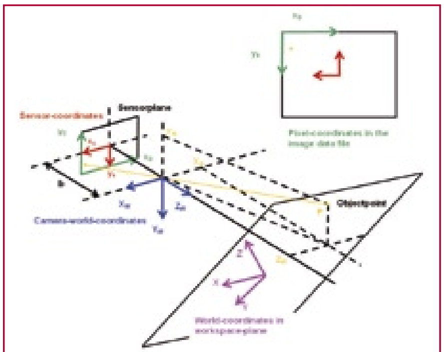

Fig. 3: Coordinate systems for the transformation of world coordinates from workspace to pixel coordinates in the image data file

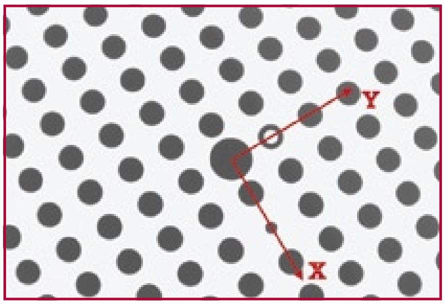

Fig. 4: Flat calibration target with X- and Y-axis of the world-coordinate system in the base-plane of the robot

Uncertainty

A simple test of the model is to try to reproduce the images of those circles on the calibration pattern, which themselves have been used for calibration. A proper figure of merit is the deviation for back-projection. The camera model with its parameters drawn from calibration is used to calculate the pixel coordinates of the reference points from their known world coordinates. These results are compared with the pixel coordinates taken as centers of mass from the real image. Usually, the data will not match perfectly, since the calibration is based on a regression of a huge set of data points with a small set of parameters. Both, the mean or the maximum deviation for back-projection are suitable measures for the uncertainty of the method. In our laboratories, we usually obtain back-projection deviations in the order of 1/10 to 2/10 of the edge length of a pixel for working distances in the order of 0.5 m with 8-bit standard analog cameras and off-the-shelf lenses. The uncertainty of the method thus is in the same order of magnitude as the uncertainty of the position of the centre of mass calculated for the reference objects from the image data. From this point of view, the method seems to be well suited. In fact, supported by a camera calibrated according to these principles, the robot shown in figure 1 precisely picks up the objects from the tray on the left and reliably places them in the moulds of the tray on the right.

The method described above, which utilizes a flat calibration pattern in a plane, by no means restricts the calibration to this plane. The camera is completely calibrated, as long as the orientation and position of the system remain untouched. In the calibration-plane, however, the three world-coordinates (X, Y, Z) may be completely and absolutely determined, since the Z-coordinate for all points in this plane is equal to zero by definition and thus known. Whenever one of the three world coordinates for an arbitrary point in the scene is known, the two remaining world coordinates may be calculated from the pixel coordinates of the corresponding point in the image. A certain point might hover 15 mm, e.g., above the plane where the calibration pattern has been placed. If the corresponding image point can be detected with good contrast, all three world coordinates can be calculated, since the Z-coordinate is already known, namely Z = + 15 mm. For the scene shown in figure 1 such a situation occurs whenever the robot is forced to pick up an object, since the top faces of these objects are not in the base plane of the robot where the calibration target usually is placed but rather pop up some millimeters – with interesting results when this particular fact is not taken into account.

The calibration procedure described in this article is based on the work of Lenz [1] and Tsai [2], published in 1987. It is remarkable that these fine methods needed 20 years to be generally acknowledged as a tool of the trade and have only in the recent years become commercially available. Meanwhile, camera calibration is a must for every notable image processing library and may well be classified as a basic method of image processing.

References

[1] Lenz, R., Informatik-Fachberichte 149: Mustererkennung 1987, S. 212-216, Springer 1987

[2] Tsai, R. Y., IEEE Journal of Robotics and Automation, Vol. RA-3 (4) 1987, 323-344

▶ 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