Camera calibration

- The calibration objective

  The objective of the camera calibration is to determine parameters defining the relationship between the reference 3D coordinate system and the camera coordinate frame, which are combined with the perspective transformation, and the parameters associated with a camera and optics.

- The camera parameters

  Generally, the parameters determined in a camera calibration procedure can be divided into two groups:

  – extrinsic parameters which are associated with the translation and rotation of the camera coordinate frame relative to the scene reference coordinate system,

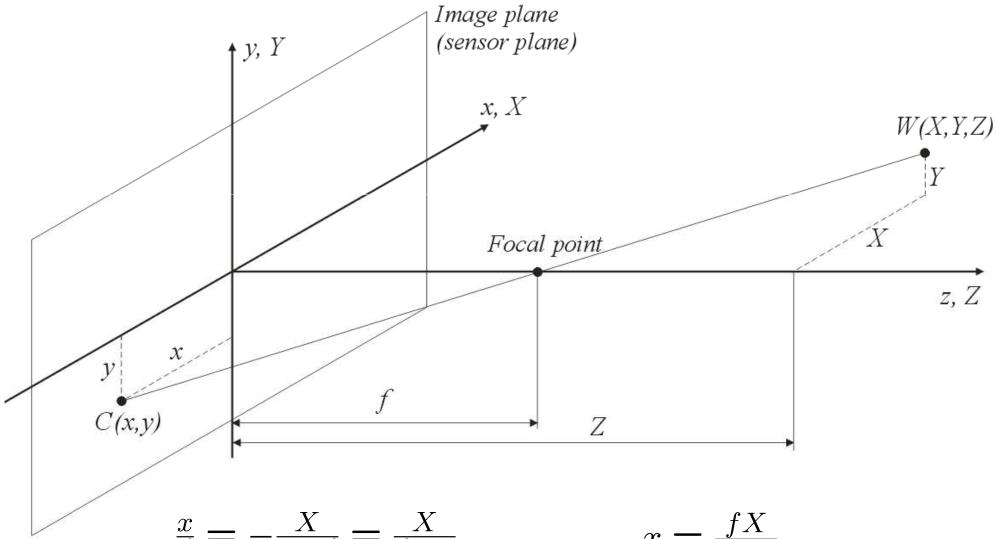
  – intrinsic parameters which describe the optical and electrical properties of the camera, the number of parameters depends on the adopted model of the optics and the camera for calibration procedure.

  In the case of intrinsic parameters, in addition to the focal length of the lens, usually radial and tangential distortions are modeled which are associated with lens optics. Additionally the real coordinates of the center of camera image (principal point) can be determined. These coordinates can be different from geometric center of the matrix sensor due to fact that centers of curvature of lens surfaces are not always strictly collinear.

Camera calibration

- The basic perspective camera model

  For a basic camera model, the points of 3D space are projected onto the plane of the camera sensor in accordance with the perspective transformation which results from physical method of image creating in the camera sensor plane.



$$\frac{x}{f} = -\frac{X}{Z-f} = \frac{X}{f-Z} \qquad x = \frac{fX}{f-Z}$$
$$\frac{y}{f} = -\frac{Y}{Z-f} = \frac{Y}{f-Z} \qquad y = \frac{fY}{f-Z}$$

Camera calibration

- The perspective transformation in terms of homogeneous coordinates

Let the coordinates of W point in Cartesian space represents the vector $\boldsymbol{w} = [X \;\; Y \;\; Z]^T$ which in homogeneous form can be written as $\boldsymbol{w}_h = [kX \;\; kY \;\; kZ \;\; k]^T$ where k is any nonzero constant. Perspective transformation matrix $\boldsymbol{P}$ can be written as

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix}$$

Homogeneous vector $\boldsymbol{c}_h$ corresponding to point C in the image coordinates can be written as follows

$$\boldsymbol{c}_h = \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix} = \boldsymbol{P}\boldsymbol{w}_h = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} = \begin{bmatrix} kX \\ kY \\ kZ \\ \frac{-kZ}{f} + k \end{bmatrix}$$

For given homogeneous image coordinates $\boldsymbol{c}_h$ Cartesian coordinates of the point C can be determined by dividing the first three elements by the fourth

$$\boldsymbol{c} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{fX}{f-Z} \\ \frac{fY}{f-Z} \\ \frac{fZ}{f-Z} \end{bmatrix}$$

Camera calibration

The inverse perspective transformation remaps image coordinates of the point to the 3D space according to the relation

$$\boldsymbol{w}_h = \boldsymbol{P}^{-1}\boldsymbol{c}_h$$

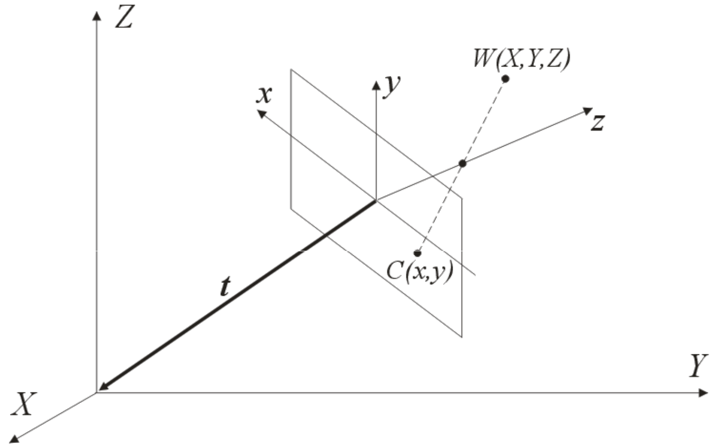
This leads to the following formulas of point coordinates in 3D space based on its image projection coordinates

$$X = \frac{x}{f}(f - Z)$$
$$Y = \frac{y}{f}(f - Z)$$

Based on a single image without some knowledge of the observed scene, for example the Z component of the points, there is not possible to fully determine a point in three-dimensional space based on its coordinates in the image coordinates. Therefore, in order to determine the full information about the location of points in 3D space, stereovision camera system can be used.

Camera calibration

- Transformations between the reference coordinate frame and the camera coordinates



$$\boldsymbol{c}_h = \boldsymbol{P}\boldsymbol{T}_T\boldsymbol{T}_R\boldsymbol{w}_h$$

Homogeneous transformation matrix $\boldsymbol{T}_T$ determines the translation between the reference coordinates and the camera frame and is expressed as

$$\boldsymbol{T}_T = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{t} \\ \boldsymbol{0} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Camera calibration

Homogeneous transformation matrix $\boldsymbol{T}_R$, describing the rotation of the reference frame to the camera frame, is composed of three elementary rotation using Euler angles. For example, the Euler angles of rotation around the fixed axes in notation Roll, Pitch, Yaw gives

$$\boldsymbol{T}_R = \begin{bmatrix} c\gamma & -s\gamma & 0 & 0 \\ s\gamma & c\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c\beta & 0 & s\beta & 0 \\ 0 & 1 & 0 & 0 \\ -s\beta & 0 & c\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c\alpha & -s\alpha & 0 \\ 0 & s\alpha & c\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} =$$

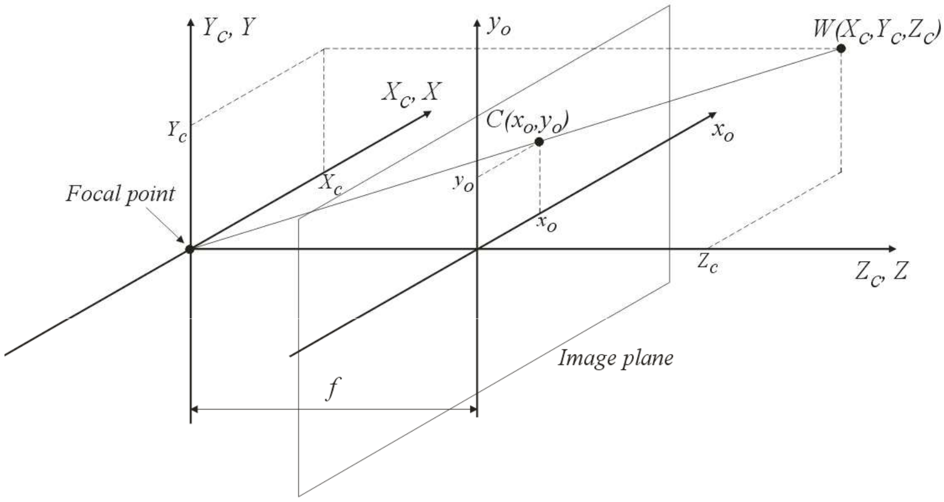
$$= \begin{bmatrix} c\beta c\gamma & s\alpha s\beta c\gamma - c\alpha s\gamma & c\alpha s\beta c\gamma + s\alpha s\gamma & 0 \\ c\beta s\gamma & s\alpha s\beta s\gamma + c\alpha c\gamma & c\alpha s\beta s\gamma - s\alpha c\gamma & 0 \\ -s\beta & s\alpha c\beta & c\alpha c\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_G^C & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix}$$

Assuming $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{T}_T\boldsymbol{T}_R$, a homogeneous transformation equation can be written as

$$\begin{bmatrix} c_{h1} \\ c_{h2} \\ c_{h3} \\ c_{h4} \end{bmatrix} = \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Camera calibration

- Pin-hole camera model



In pin-hole camera model the projection from 3D space to the image plane can be represented as perspective transformation with the additional rotation of the image plane by 180 degree. In this model the camera coordinate frame (X_c, Y_c, Z_c) has the origin at the focal point, and the Z_c axis coincides with the optical axis of the camera. The image coordinate frame is placed in the front of focal point at the focal distance f instead of $-f$. Such a theoretical model of the camera, which in practice does not exist, is simpler to analyze.

Camera calibration

For pin-hole camera model the following relationships between a point $W(X_c, Y_c, Z_c)$ in the camera coordinates (X_c, Y_c, Z_c) and its projection $C(x_o, y_o)$ on the image plane in the image coordinates (x_o, y_o) can be written as

$$x_o = f\frac{X_c}{Z_c}$$
$$y_o = f\frac{Y_c}{Z_c}$$

The image coordinates of the point are the normalized coordinates only with respect to the coordinate Z_c of the point in the camera frame. Such representation is more convenient in comparison to the classical perspective transformation.

In situations where the reference system is not identical with the camera frame, before a projection of the point onto the image plane, the rotation and translation transformation should be performed from the reference system to the camera coordinate frame

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \boldsymbol{R} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \boldsymbol{t}$$

where the matrix $\boldsymbol{R}$ determines the rotation and vector $\boldsymbol{t}$ translation, which represent extrinsic parameters of the camera.

Camera calibration

- Optical distortion in the camera model
  - J. Heikkila, O. Silven, *A Four-step Camera Calibration Procedure with Implicit Image Correction*, Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1106-1112
  - J.Y. Bouguet, *Camera Calibration Toolbox for Matlab*, http://www.vision.caltech.edu/bouguetj/calib_doc/

Distortions of the lens is usually modeled as radial, tangent and linear distortion. Radial distortion causes the actual image point to be displaced radially in the image plane and has a decisive influence on the measurement accuracy. The radial distortion can be approximated using the following expression

$$\begin{bmatrix} \delta_x^r \\ \delta_y^r \end{bmatrix} = \begin{bmatrix} x_o(k_1 r^2 + k_2 r^4 + \ldots) \\ y_o(k_1 r^2 + k_2 r^4 + \ldots) \end{bmatrix}$$

where $k_1, k_2, \ldots$ are the coefficients of radial distortion, $r = \sqrt{x_o^2 + y_o^2}$.

Typically, one, two or three coefficients are enough to compensate the radial distortion.

Camera calibration

Tangent distortion result from the fact that the centers of curvature of the lens surface are not always strictly collinear. This decentering distortion has both a tangential and radial component and is modeled by the following expression

$$\begin{bmatrix} \delta_x^t \\ \delta_y^t \end{bmatrix} = \begin{bmatrix} 2p_1 x_o y_o + p_2(r^2 + 2x_o^2) \\ p_1(r^2 + 2y_o^2) + 2p_2 x_o y_o \end{bmatrix}$$

where p_1 and p_2 are tangential distortion coefficients.

Linear distortions can be described by a single curvature ratio α_c. This term is relevant if the image axes are not orthogonal. The magnitude of linear distortion is related to quality and precision of the sensor used in the camera.

Camera calibration

- Intrinsic parameters of the camera model with distortion coefficients

  Taking into account the distortion in the camera model, one can specify the following intrinsic camera parameters for the calibration procedure:

  – two-element vector of the focal length $\boldsymbol{f}_c$, whose elements are linear relationship of focal length and scale factors resulting from recalculation of the pixels size in both axes from metric units to image coordinates (focal length is expressed in pixels),

  – two-element vector $\boldsymbol{c}_c$ containing the real coordinates of the center point of projection (principal point) in the image,

  – skew coefficient α_c of linear distortion, which defines the skew angle between the axes of pixels in the horizontal and vertical direction in the sensor matrix,

  – five-element distortion vector $\boldsymbol{k}_c$ containing the coefficients of radial and tangential distortion, where

  $\boldsymbol{k}_c = [k_1 \ k_2 \ p_1 \ p_2 \ k_3]^T$

Camera calibration

- The real coordinates of the image coordinate system including the intrinsic parameters

  The normalized coordinates of the point in the image coordinates

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} \frac{X_c}{Z_c} \\ \frac{Y_c}{Z_c} \end{bmatrix}$$

For the introduced internal parameters and normalized image coordinates the real coordinates of the point with regard to the equations defining the radial and tangential distortion can be written as

$$\boldsymbol{x}_d = \begin{bmatrix} x_d(1) \\ x_d(2) \end{bmatrix} = (1+k_c(1)r^2+k_c(2)r^4+k_c(5)r^6) \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \boldsymbol{d}_x$$

Vector $\boldsymbol{d}_x$ consist of tangent distortion and is written as

$$\boldsymbol{d}_x = \begin{bmatrix} 2k_c(3)x_n y_n + k_c(4)(r^2 + 2x_n^2) \\ k_c(3)(r^2 + 2y_n^2) + 2k_c(4)x_n y_n \end{bmatrix}$$

Camera calibration

Taking into account the vector of the camera focal length, the real position of the center of the image and the skew coefficient, the coordinates of the point in the image coordinate frame can be represented as

$$x_o = f_c(1)(x_d(1) + \alpha_c x_d(2)) + c_c(1)$$
$$y_o = f_c(2)x_d(2) + c_c(2)$$

The above relations are often presented in the form of a matrix equation

$$\begin{bmatrix} x_o \\ y_o \\ 1 \end{bmatrix} = \boldsymbol{K}_k \begin{bmatrix} x_d(1) \\ x_d(2) \\ 1 \end{bmatrix}$$

where matrix $\boldsymbol{K}_k$ is so called camera matrix

$$\boldsymbol{K}_k = \begin{bmatrix} f_c(1) & \alpha_c f_c(1) & c_c(1) \\ 0 & f_c(2) & c_c(2) \\ 0 & 0 & 1 \end{bmatrix}$$

Elements of the matrix $\boldsymbol{K}_k$ and distortion vector $\boldsymbol{k}_c$ should be determined in camera calibration procedure together with the extrinsic parameters.

Camera calibration

- Camera calibration procedures

Most of the calibrating methods solves the calibration problem in two steps, with reference to the Tsai[1] method, which was the first proposition of two-step calibration procedure. In the first step, using linear methods, an approximate solution is obtained and next the result is used in the second stage as the initial condition to the non-linear optimization. Thanks to decomposition of the calibration problem in two steps the acceleration of calculation is obtained. More over the calibration accuracy is better in comparison to conventional methods taking into account the distortions.

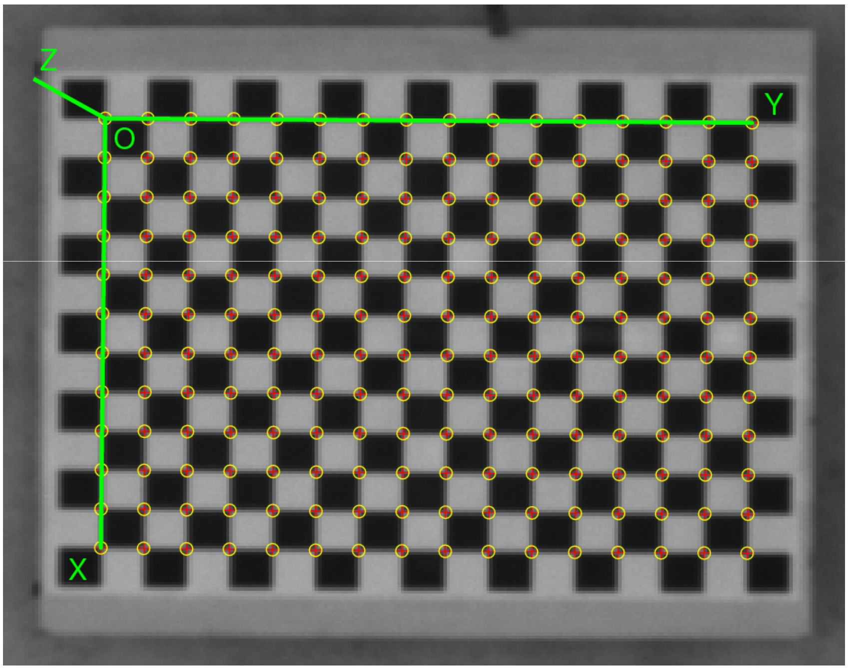
Any calibration method requires knowledge of the series of points in 3D space and their projections coordinates on the image plane. The calibration methods can be typically divided into three categories according to manner of calibration points determining:

- the first is based on the three-dimensional reference object whose shape, dimensions and location are known with high precision. A variation of this approach is to use a planar pattern with a known change of the pattern position when determining the calibration points,
- the second method comes from the observation of plane pattern placed in different positions relative to the camera. However, there is no need to know the pattern position in 3D space but only the relationship between calibration points on the pattern,
- the third category is determined as a self-calibration, which does not use any object. This method is based on corresponding points in the images acquired during the movement of the camera and observing a static scene.

[1] R.Y. Tsai, An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, pp. 364-374, 1986.

Camera calibration

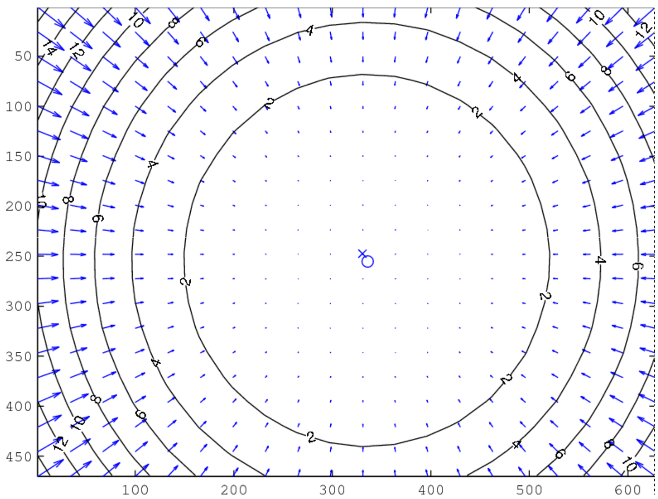
- The pattern for the calibration procedure with coordinate frame and the determined calibration points



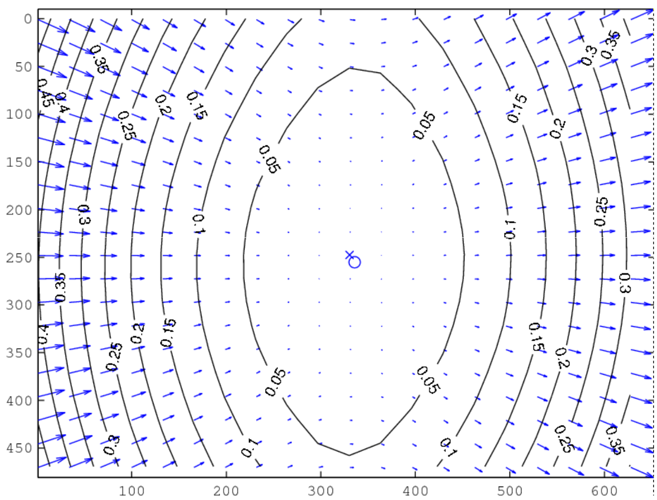
Camera calibration

- The distribution of radial and tangential distortion

camera: VISTEK SVS 084MSCL , lens: PENTAX H612A



Radial distortion



Tangential distortion

Camera calibration

- Inverse problem

The solution of the inverse problem allows for correction of distortion in the image, points reprojection from image coordinates and measurements in the metric units in reference coordinate system with the known transformations associated with the rotation and translation. Inverse problem consists in determination of the vector of normalized coordinates $[x_n \;\; y_n]^T$ based on the pixel image coordinates (x_o, y_o), using knowledge of the camera model and the parameters.

Due to the nonlinearity of the transformation associated with distortion in the camera model, there is no possibility to propose a general algorithm solving the inverse problem!

For the proposed model with radial and tangential distortion the normalized coordinates of the point can be determined by the following procedure

1. For the pixel image coordinates (x_o, y_o) the initial normalized vector $[x_n \;\; y_n]^T = \boldsymbol{x}_d$ can be calculated as

$$x_d(1) = (x_o - c_c(1))\,/f_c(1)$$
$$x_d(2) = (y_o - c_c(2))\,/f_c(2)$$

Camera calibration

2. calculate the tangent distortion vector $\boldsymbol{d}_x$ and radial distortion factor k_{radial} as

$$\boldsymbol{d}_x = \begin{bmatrix} 2k_c(3)x_n y_n + k_c(4)(r^2 + 2x_n^2) \\ k_c(3)(r^2 + 2y_n^2) + 2k_c(4)x_n y_n \end{bmatrix} \qquad r^2 = x_n^2 + y_n^2$$

$$k_{radial} = 1 + k_c(1)r^2 + k_c(2)r^4 + k_c(5)r^6$$

3. calculate the new normalized coordinates

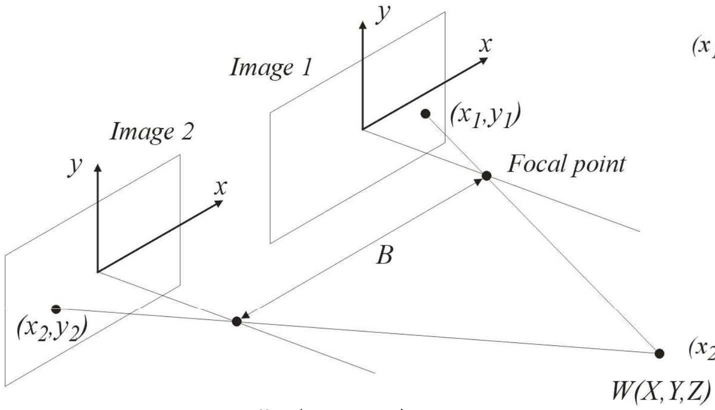
$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = (\boldsymbol{x}_d - \boldsymbol{d}_x)\,/k_{radial}, \qquad k_{radial} \neq 0$$

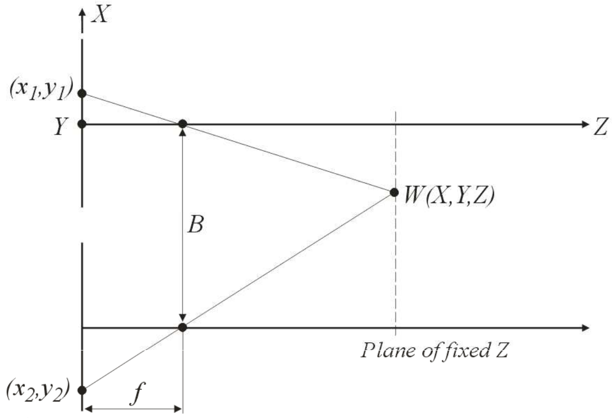
4. repeat steps 2-3 fixed number of times or test whether matching error is less than assumed, the tracking error can be calculated as

$$\triangle \boldsymbol{D} = \boldsymbol{x}_d - \left(k_{radial} \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \boldsymbol{d}_x \right)$$

Stereovision

- Stereovision with two cameras





$$X_1 = \frac{x_1}{f}(f - Z)$$

$$X_2 = \frac{x_2}{f}(f - Z)$$

$$X_2 = X_1 + B$$

$$Z = f - \frac{fB}{x_2 - x_1}$$

$x_2 - x_1$ - disparity

$$X = \frac{x_1}{f}(f - Z)$$

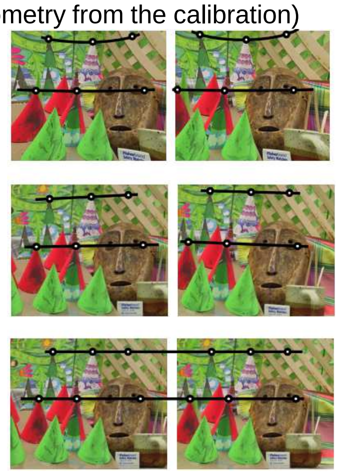
$$Y = \frac{y_1}{f}(f - Z)$$

$$Z = f - \frac{fB}{x_2 - x_1}$$

Stereovision

- Stereovision with two cameras

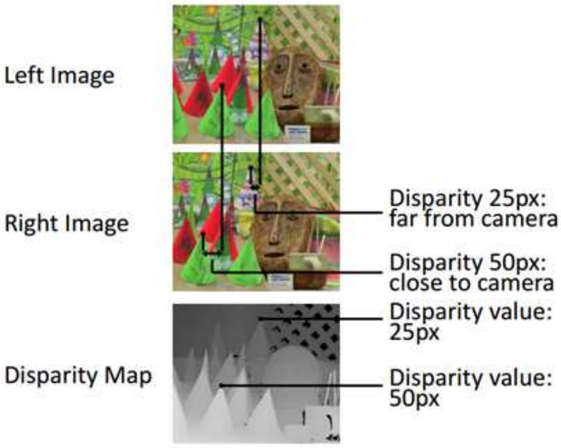
1. Image rectification (use of the cameras geometry from the calibration)



2. Image segmentation and stereo matching (proces of finding corresponding pixels or regions in the images)

Stereovision

3. Disparity map



4. Reprojection from disparity map to 3D space