Module 7 : Robot vision I

Lecture 26 : The Imaging Transformation

Objectives

In this course you will learn the following

- Relationship of real world with oamera image.
- Transfer of image.

Introduction

So far we have seen how video images are acquired and how we can represent them in a digital form. In this lecture, we consider the relationship between the real world in front of the camera and the image obtained in the camera. We will make use of two coordinate frames of reference in studying this relationship. We will assume the *World* coordinate frame to be the reference for the locations of the objects of interest in the real world. We will use capital letters to denote quantities in this coordinate frame of reference. We will also make use of a coordinate frame fixed to the image plate of the camera (in the focal plane) and call it the image coordinate frame. We will use small letters to denote quantities in this coordinate frame of reference.

In robot vision applications in particular, we would be interested in knowing where in the world coordinate frame does a point lie given its image coordinates. As can be seen easily, it is not possible to accomplish this task without additional information about the point of interest. The reason is that the real word is three dimensional and image is only two-dimensional. Thus an image does not have sufficient information to reconstruct the real world. This problem can be sorted out by using additional camera (stereo imaging) or information from another sensor, such as a range sensor, or information about an object in view, such as its size, etc.

Consider a point P in world coordinates P(X, Y, Z). This point will have image C in Image plane as shown below through a camera lens having focal length λ . The Point C will have coordinates C(x, y, z). Let us take the two coordinate frames of reference to be coincident for now.

Figure 1: The geometry of image formation in a camera system

From the properties of similar triangles, we have

$$x / \lambda = -X / (Z - \lambda) = X / (\lambda - Z) \qquad (1)$$

$$y / \lambda = -X / (Z - \lambda) = Y / (\lambda - Z) \qquad (2)$$

We can make use of homogenous coordinates that were used in analyzing robot kinematics to write these relationship in a more compact way as,

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

Recall that the physical 3D coordinates are obtained from the 4x1 homogeneous coordinate vector by dividing the first three elements of the vector by the fourth one and eliminating the fourth element. The 4x4 matrix, P is known as the *imaging transformation* or the *perspective transformation* matrix. Given the homogenous world coordinates W h and the homogeneous image coordinates $c_h$can be expressed by above equation.

$$c_h = P\, W_h \tag{5}$$

Contd...

We will be more interested in the inverse problem of finding the world coordinates given the image coordinates. We can find the inverse imaging transformation matrix by inverting the P matrix. Thus;

$$W_h = P^{-1}\, c_h \tag{6}$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/\lambda & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 + z/\lambda \end{bmatrix} \tag{7}$$

In the normal vector form this gives us,s

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \lambda x/(\lambda + z) \\ \lambda y/(\lambda + z) \\ \lambda z/(\lambda + z) \end{bmatrix} \tag{8}$$

The problem here is that z is meaningless as the image is only two dimensional and z=0 for the image plane. Let us substitute z in the first two equations in the above in terms of Z as obtained from the third equation.

$$z = \lambda Z/(\lambda - Z) \tag{9}$$

The first two equations reduce to,

$$\begin{aligned} X &= x\,(1 - Z/\lambda) \\ Y &= y\,(1 - Z/\lambda) \end{aligned} \tag{10}$$

Which implies that the world coordinates of the point of interest in the real world can be found provided its distance from the image plane is known. In stereo imaging, the second camera is able to give us this information. Alternately range sensing can be used to estimate this distance. Let us see how this done in stereo imaging.

Contd...

Example:

Consider two identical cameras, distance B apart, with the optical axes parallel and image planes in the same plane. Given a point of interest, W in the world coordinate frame of reference, with the corresponding images at (x1 y1) and (x2 y2), determine the location of W in the world coordinate frame.

The focal length of the lens is λ .

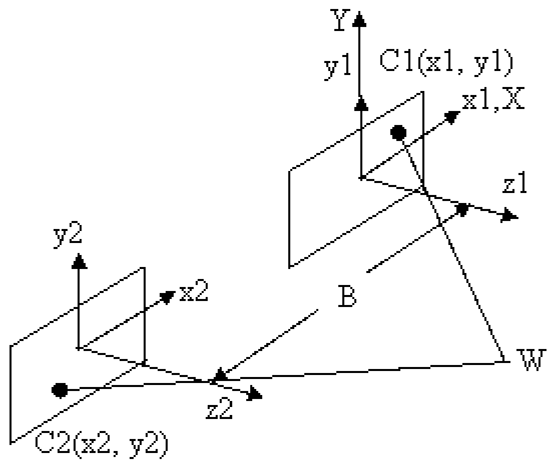

Figure 2: Stereo vision system

Solution:

We will assume that the x axes of the two image coordinate frames of reference are aligned to each other. Let world coordinate system be coincident with first camera coordinate system.

As we have seen earlier, from similar triangles in the figure,

$$X_1 = \frac{x_1}{\lambda}(\lambda - Z_1) \tag{10}$$

$$X_2 = x_1 + B = \frac{x_2}{\lambda}(\lambda - z_2) \tag{11}$$

but $z_1 = z_2 = z$. From the above equations, we have

$$z = \lambda - \frac{\lambda B}{(x_2 - x_1)} \tag{12}$$

The point of interest in world coordinate frame is therefore,

$$\begin{aligned} X_1 &= \frac{x_1 B}{(x_2 - x_1)} \\ Y_1 &= \frac{y_1 B}{(x_2 - x_1)} \end{aligned} \tag{13}$$

Thus, X, Y and Z can be found out from the images of two cameras

Closure:

We have seen that locations of points in the real world are related to the corresponding points in the camera image through a transformation in the homogenous coordinates. However, this transformation is many to one since the image z coordinate is always zero. This is a reflection of the fact that all points on the line passing through the lens centre map to the same image points. For this reason, the imaging transformation is also known as perspective transformation. Therefore the inverse problem can only be solved if additional information is available regarding the distance of the point of interest from the image plane.

We saw that stereo imaging provides us one way to obtain the information about the distance. A range sensor can also provide this information directly. Another way to obtain the distance information is possible if the actual distance between two points in the real world were known and both points are seen in the image.

Recap

In this course you have learnt the following

- Relationship of real world with oamera image.

- Transfer of image.

Congratulations, you have finished Lecture 26. To view the next lecture select it from the left hand side menu of the page.