Assume \( z_0 \gg f \) and \( f = 1 \). \( x_0, y_0, z_0 \) measured relative to 
G FRAME, \( x_i, y_i \) are points in \( i \) (image) FRAME.

\[
X_0 = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} f \cdot x_0/z_0 \\ f \cdot y_0/z_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} f x_0/z_0 \\ f y_0/z_0 \\ z_0 \end{bmatrix}
\]  

(5.11)

CASE: Vector \( \vec{X}_i = x_i' - x_i \) as a "flow vector" or perturbation, or disparity vector.

- all pts. moving represent a
  - flow or disparity field

No unique. If motion along

\( x_i' \) passing through anti projection \( t = 0 \)
Normally, we know the size of the image plane in the camera. Furthermore, we know the size of the frame grabber frame in pixels. A common image plane used is 8.8 mm by 6.6 mm. Hence, if the number of pixels on a row is 512, the number of pixels per mm of the image plane is:

\[
\text{Pixels per mm}_i = \frac{512 \text{ pixels}}{8.8 \text{ mm}} = 58.182
\]
or
\[
\text{mm per pixel}_i = \frac{8.8 \text{ mm}}{512 \text{ pixels}} = 0.01718745
\]
or
\[
\text{inches per pixel}_i = \text{mm per pixel}_i \times (1 \text{ inch} / 25.4 \text{ mm}) = 6.76671 \times 10^{-4}
\]  

(5-11a)

All of the above apply only to the 8.8 x 6.6 mm image plane.

Referring to Figure 5.8 on page 41-1 and using geometry we can write:

\[
\frac{y_i (a \text{ horizontal point in image plane})}{f} = \frac{y_0 (\text{same horizontal point in Global, Real World frame})}{z_0 + f}
\]

(5-11b)

Typically, \(z_0\), range, >> \(f\); hence, we write (5-11b) as:

\[
\frac{y_i}{f} = \frac{y_0}{z_0}
\]

(5-11c)

or

\[
f = \left( \frac{y_i}{y_0} \right) z_0 = M \times z_0
\]

(5-11d)

where

\[
\left( \frac{y_i}{y_0} \right) = M = \frac{\text{Object Size in Frame(in, ft, m, etc.)}}{\text{Object Size in Real World(in, ft, m, etc.)}}
\]

(5-11e)

Substituting (5-11d) into (5-11) on page 41-1, we have:

\[
\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} (M \times z_0)x_0 \\ z_0 \\ (M \times z_0)y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} M \times x_0 \\ M \times y_0 \end{bmatrix}
\]

(5-11f)

where \(x_i, y_i\) are in in, ft, etc.. Normally we obtain \(x_i, y_i\) in pixels. We convert \(x_i, y_i\) to units of measure in the Image Frame by multiplying \(x_i, y_i\) by a conversion factor that is in Inches(Image Frame)_per_pixel(Image Frame). Next we find \(M\), the magnitude factor, in (5-11c). Say that the object is 24 inches wide. Then we can calculate \(M\) as:

\[
M = \frac{\text{Width of Object in pixels in the Image Plane}}{\text{24 inches in Real World}}
\]

(5-11g)

or

\[
M = \frac{f(\text{inches})}{z_0(\text{inches in Global Frame})}
\]

(5-11h)
Now we can find $x_o$ and $y_o$, points on the object in Global Real World dimensions. From (5-11f) we can write:

$$x_o (\text{inches in Global Frame}) = \frac{x_i (\text{inches in Image Frame})}{M}$$

$$y_o (\text{inches in Global Frame}) = \frac{y_i (\text{inches in Image Frame})}{M}$$

$$z_o (\text{inches in Global Frame}) = \frac{f (\text{inches})}{M}$$

(5-11i)
Assume object point motion is constant velocity (and common to all object points) with respective velocities in the $x$, $y$, and $z$ directions given by $u$, $v$, and $w$. The time-varying evolution of a specific object point is

$$
\begin{bmatrix}
    x_o(t) \\
    y_o(t) \\
    z_o(t)
\end{bmatrix}
= 
\begin{bmatrix}
    x_o + ut \\
    y_o + vt \\
    z_o + wt
\end{bmatrix}
= \text{MEASURED IN GLOBAL FRAME} 
\tag{5-12}
$$

where $(x_o, y_o, z_o)$ is the location of the specific object point at $t = 0$. Thus, as a function of $t$, the motion of this object point in 3-D space generates a line (see the problems). Using Eq. 5-11, the corresponding image plane motion is

$$
\begin{bmatrix}
    x_i(t) \\
    y_i(t)
\end{bmatrix}
= 
\begin{bmatrix}
    (x_o + ut)/(z_o + wt) \\
    (y_o + vt)/(z_o + wt)
\end{bmatrix}
\tag{5-13}
$$

THE FOCUS OF EXPANSION (FOE)

Assume that the motion of the point is toward the image plane (i.e., $w = dz/\text{dt} < 0$). Due to the foreshortening property of the $p-p$ transform, as $t \to -\infty$ the motion of this object point appears to emanate from a fixed point in the image.
plane, given by

\[ \mathbf{x}(t) = \begin{bmatrix} u/v \\ v/w \end{bmatrix} = \mathbf{x}^* \text{ (FOE)} \]

(5.14)

This point in the image plane, denoted as \( \mathbf{x}^* \), is referred to as the focus of expansion, or FOE. With this model, the motion of all object points (notice the FOE is independent of \( (x, y, z) \), as seen in the image plane, appears to emanate from (or toward) \( \mathbf{x}^* \) the FOE. Figure 5.10 illustrates the significance of the FOE for a simple case. Figure 5.11 illustrates the concept with real imagery.

Several approaches to calculate the FOE exist, based on exploitation of the fact that for constant velocity object motion all image plane flow vectors intersect at the FOE. An obvious approach involves plotting of the extracted flow vectors and extrapolation backwards (or forwards) to find this intersection. An alternate approach is shown in Jain (1983).

In closing, we note that if the object point motion is not simple constant velocity translation (e.g., the object is rotating), a FOE may not exist (see the problems).
The case of multiple translating objects is also of interest: in this case each object has a FOE.

The Time-to-Adjacency Equation and Depth from Motion

It is useful to define the time-varying distance, $D(t)$, in the image plane, of an image point from the FOE. This is given as

$$D(t) = \|x_t - e\|$$

(5.15)