Let a line be defined by two 2D pts in homogeneous coordinates: $A = (x, y, t)$ & $B = (0, 0, w)$.

:: The line passes through $(\frac{x}{t}, \frac{y}{t})$ & $(\frac{0}{w}, \frac{w}{w})$.

Let $m$ & $c$ be the slope & offset of the eqn of this line. Hence,

$$\frac{w}{w} = m \cdot \frac{w}{w} + c$$

$$\frac{y}{t} = m \cdot \frac{x}{t} + c$$

:: $$\left(\frac{w}{w} - \frac{y}{t}\right) = m \left(\frac{w}{w} - \frac{x}{t}\right)$$

$$\frac{tw - yw}{wt} = m \left(\frac{tw - xw}{wt}\right)$$

:: $m = \frac{tw - yw}{tw - xw}$

:: $c = \frac{w}{w} - \frac{tw - yw}{tw - xw} \cdot \frac{w}{w}$

$$= \frac{tw - xw - tw + yw}{w(tw - xw)}$$
\[
\frac{y_t (y_u - x_u)}{x_t (t_u - x_u)}.
\]

Note that any pt. \( P \) (3x1 vector) should satisfy the eq. \[
\begin{bmatrix}
    y_u - t_u \\
    t_u - x_u \\
    x_u - y_u
\end{bmatrix} P = 0 \] to lie on the line. \( \vdots \) This matrix signifies the line eqn. \( \vdots \) Note \[
\begin{bmatrix}
    y_u - t_u \\
    t_u - x_u \\
    x_u - y_u
\end{bmatrix} \times (x, y, t) = (u, v, w) \times \begin{bmatrix}
    x \\
    y \\
    t
\end{bmatrix} = B \times A
\]

Now \[
\begin{bmatrix}
    y_u - t_u \\
    t_u - x_u \\
    x_u - y_u
\end{bmatrix} \times \begin{bmatrix}
    0 \\
    w \\
    -u
\end{bmatrix} = \begin{bmatrix}
    -w & 0 & u \\
    v & -u & 0
\end{bmatrix} \times \begin{bmatrix}
    B \\
    x
\end{bmatrix}
\]

where \( B_m \) is the special matrix formed by the 3D homogeneous coordinates of a 2D pt.

Also \[
\begin{bmatrix}
    y_u - t_u \\
    t_u - x_u \\
    x_u - y_u
\end{bmatrix} = A^T [B]_x^T
\]
Hence if $A$ is a pt. on a line $AB$, the line eqn. is given by

$[B]_x^T A$ or $A^T [B]_x^T$.

Now if $P$ satisfies this eqn. then

$12^T ([B]_x A) = (A^T [B]_x^T) P = 0$

Note $\det ([B]_x) = 0$, all $2 \times 2$ submatrices have $\det = 0$. \( \therefore \) Rank 2 matrix.

**EPIPOLAR GEOMETRY**

It deals with several constraints and invariants when considering a pair of cameras. This helps in several problems like depth reconstruction & motion estimation.
Consider two cameras, $O_1$ & $O_2$ are the COP of the two cameras. $O_1O_2$ is often called the baseline, especially considering $O_1$ & $O_2$ as a stereo pair.

1. Consider a 3D pt. $P$. Let $p$ be its image in camera $O_1$ & $p'$ be its image in $O_2$. $O_1O_2$ defines a plane. Note that as $P$ changes, this plane changes but rotates about $O_1O_2$. This defines a pencil of planes rooted at $O_1O_2$.

2. Note that the image of any pt. on the ray $QP$ forms a line $l'$.
3. The line joining $O_1O_2$ intersects image plane $I$ of $O_1$ & $O_2$ at pts. $E_1$ & $E_2$ respectively. These are called the epipoles of $O_1$ & $O_2$ respectively.

4. The lines $E_1p = l$, $E_2p' = l'$ are the epipolar lines. Note that as the plane $P$ of $O_1O_2$ changes, since $O_1$ & $O_2$ are fixed, the epipoles don't change, hence all the epipolar lines pass through the epipole of the image.

Why is this important?
Assume calibrated camera & stereo depth reconstruction. First, it reduces the search space for correspondence. If we detect feature $p$ in $C_1$, then we need to search for its correspondence on the line $E_2p'$ instead of the whole image. Thus, reduces the search.
Space from a 2D plane to 1D line.

**Fundamental Matrix**

If \( P = (x, y, t) \) & \( q = (u, v, w) \),

The line \( l \) is given by the matrix

\[
l = \begin{pmatrix} 0 & w - u & x \\ -v & 0 & y \\ -u & -w & t \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = Lp
\]

Note that \( l' \) & \( l \) are coplanar. So there exists a 2D affine transformation to map one to the other. Let this be denoted by a \( 3 \times 3 \) matrix \( A \).
\[ l' = A l \]
\[ l' = A L \begin{pmatrix} m \\ y \\ t \end{pmatrix} \]
\[ = F \begin{pmatrix} m \\ y \\ t \end{pmatrix} = Fp \]

Now since \( p' = (n', y', t') \) lies on this line we will get
\[ p'^T Fp = 0 \]

\[ \therefore p' \text{ satisfies the line} \]

\[ Eq = \]

\( F \) is called the fundamental matrix.

\[ \therefore \text{Point } p \text{ defines a line } l' \text{ on which } p' \text{ lies.} \]
Estimating The Fundamental Matrix

\[
\begin{bmatrix}
x' \\
y' \\
1
\end{bmatrix}
\begin{bmatrix}
f_1 & f_2 & f_3 \\
f_4 & f_5 & f_6 \\
f_7 & f_8 & f_9
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix} = 0
\]

\[
\therefore x'n'f_1 + xy'n'f_2 + xf'_3 + yn'f_4 + y'y'n'f_5 + y'y'n'f_6 + y'n'f_7 + y'n'f_8 + f_9 = 0
\]

\therefore With multiple correspondences you can solve for \( f \). What is the minimum number of points?

\( \| f \| = 1 \) since correct up to a scale factor. \( \therefore \) Need at least eight pts.
Properties
1. \( F \) is a rank 2 matrix with 7 degrees of freedom
2. \( P' F P = 0 \) \( \Rightarrow \) 2 rotations
   \( \Rightarrow \) 2 translations
   \( \Rightarrow \) 3 params \( \mathbf{R} \), \( \mathbf{t} \).
3. \( \lambda' = FP \)
   \( \lambda = F P' \)
4. \( F e_1 = 0 \)
   \( F^T e_2 = 0 \)

Difference from homography

.: you plug in a pt \( \mathbf{p} \) & get a corresponding pt. \( \mathbf{p}' \) in the second image.

But in this case, the constraint
does not allow you to find the pt. If you are given p, you can find Fp. But this is not p'. p' is a pt. which falls on the line defined by Fp and hence you have to search that line. (Normalized Coordinates provides a well-conditioned system).

Say Fp is \[
\begin{bmatrix}
  \mathbf{f} \\
  \mathbf{f}_1 \\
  \mathbf{f}_2 \\
  \mathbf{f}_3
\end{bmatrix}
\]

This is a line whose slope is given by \( \frac{\mathbf{f}_1}{\mathbf{f}_3} \) and offset by \( \frac{\mathbf{f}_2}{\mathbf{f}_3} \).
Let us assume that $C_1$ and $C_2$ are the two camera calibration matrices.

$$C_2 \ 0_1 = e_2$$

**Let the 3D point $P$ in**

$$P = C_1 \ p$$

$$\implies P = C_1^+ \ p \quad \text{where } C_1^+ \text{ is the pseudo inverse.}$$

$$\implies P' = C_2 \ C_1^+ \ p$$

**The line $e_2 \ P'$ is given by**

$$C_2 \ 0_1 \times C_2 \ C_1^+ \ p$$

$$\implies e_2 \times C_2 \ C_1^+ \ p$$

$$\implies [e_2 ]_x \ C_2 \ C_1^+ \ p$$

$$= Fp$$
Now note that
\[
P = \begin{bmatrix} e_2 \end{bmatrix}^T \begin{bmatrix} C_2 & C_1^+ \end{bmatrix}_{3 \times 3} \begin{bmatrix} 3 \times 3 \ 3 \times 4 \ 4 \times 3 \end{bmatrix}
\]

\[\therefore e_2 C_1^+ \text{ is a } 3 \times 3 \text{ matrix. This is exactly the homography via the plane } \Pi_1 \text{ defined by } P, \quad 0_1, \quad \text{and } 0_2.\]

\[\text{rank } 2 \leftrightarrow \text{rank } 3 \quad \text{rank } 2 \quad \text{rank } 3\]

\[\therefore P = \begin{bmatrix} e_2 \end{bmatrix}^T H_{\Pi_1} \]

Let us consider a calibrated stereo rig. \( C_1 = k_1 [I | 0] \), \( C_2 = k_2 [R_1 | t] = [k_2 R_1 | k_2 t] \), \( C_1^+ = \begin{bmatrix} k_1^{-1} \\ 0 \end{bmatrix} \), \( C_2 C_1^+ = k_2 R_1 k_1^{-1} \)

Since \( 0_1 = (0, 0, 0) \), \( \therefore C_2 0_1 = k_2 t \)
\[ F = \begin{bmatrix} e_2 \end{bmatrix} \begin{bmatrix} c_2 & c_1^+ \\ \end{bmatrix} \]

\[ = \begin{bmatrix} e_2 \end{bmatrix} x
\]

\[ = \begin{bmatrix} e_2 \end{bmatrix} k_2 R_1 k_1^{-1}
\]

\[ = \begin{bmatrix} c_2 o_1 \end{bmatrix} x
\]

\[ k_2 R_1 k_1^{-1} = \begin{bmatrix} k_2 + \end{bmatrix} x
\]

If given calibrated camera, you can easily find the fundamental matrix.

How does it help?

Say two camera positions of the same camera obtained after a pure translation.

\[ C_1 = k \begin{bmatrix} I & 0 \end{bmatrix} \]

\[ C_2 = k \begin{bmatrix} I & t \end{bmatrix} \]

\[ F = \begin{bmatrix} e_2 \end{bmatrix} K K^{-1}
\]

\[ = \begin{bmatrix} e_2 \end{bmatrix} x
\]

\[ K K^{-1} \]

\[ R_I = I = R_2 \]

\[ k_1 = k_2 = K \]
If the t is parallel to x axis, then epipole is on x axis at infinity.

\[ e_2 = (0) \]

\[ [e_2]_x = P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \]

\[ [x', y', 1] \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} n' \\ y' \\ 1 \end{bmatrix} = 0 \]

\[ a \begin{bmatrix} 0 & 1 & -y' \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} n' \\ y' \\ 1 \end{bmatrix} = 0 \]

\[ y - y' = 0 \]

\[ y = y' \]

\[ \text{Epipolar lines are raster lines (lines parallel to x axis) in each image.} \]

\[ \text{Very easy to find correspondences.} \]
Essential Matrix

If you have a camera whose $K$ is identity, then

$$C = (R | t)^T$$

This is called a normalized camera. This is achieved when the camera coordinates are normalized & hence the name.

Now if we consider two normalized cameras, then their fundamental matrix is called an essential matrix.

Two normalized cameras satisfy,
\[ \hat{p}' E \hat{p} = 0 \]

where \( \hat{p} \) indicates normalized camera coordinates, \( E \) is essential matrix.

For general cameras, the essential matrix is given by
\[ E = k_2^T F K_1 \]  \[ \text{[Proof not included]} \]

**Estimating Essential Matrices**

Apply the same method as estimating fundamental matrix but using normalized camera coordinates. This assumes that there are normalized cameras.

It can be shown that for normalized cameras
\[ E = R [t]_x \]
where \( R \) & \( t \) are the transformation
Required to the image planes for two cameras parallel when \( y = y' \) – correspondences lie on raster lines.

So if we can estimate \( R \& t \), then we can apply it to an image & get the correspondences easily searchable on raster lines. This is called Rectification.

(1) 

(2)
How do you estimate \( R \) & \( \Sigma \)?

\[ \Sigma = U \Sigma V^T \]  

Proof not included.

\( U \) & \( V \) are 3x3 orthogonal matrices.

\[ \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix} \]

By internal constraints on \( E \). Define orthonormal

\[ W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

\[ W^{-1} = W^T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

\[ E = VW^T \]

\[ R = UW^{-1}V^T \]  

\[ \text{Proof:} \quad REJ_x = UW^{-1}V^TVWEV^T = UW^TV^T \\

\text{(1)} \]

\[ RJ_x = VW^T \]
\[ R = U \Sigma V^T \]
\[ = U \Sigma V^T = \mathbf{e} \]

Actually a solution:

\[ R = U W V^T \]
\[ \sigma [t]_x = \sqrt{w^{-1} \Sigma V^T} \]

will also work:

\[ R = U W V \]
\[ \sigma [t]_x = \sqrt{V^T w^{-1} \Sigma V^T} \]

\[ w \]
\[ R = U W^{-1} V \]
\[ \sigma [t]_x = \sqrt{V^T w \Sigma V^T} \]

\[ \ldots \text{four solutions to this problem.} \]
Only one of the 4 results are possible in practice. Often will generate 3D pts. behind 1st, 2nd or both cameras.

Rectification is very common in stereo matching procedure to reduce complexity of correspondence search.

After rectification images are very close to the two eyes. i.e. Two cameras with parallel image planes separated by a translation.
Let us investigate the eye situation.

\[ \therefore \beta + \alpha < \theta \]
1. A point inside horopter moves outward.
2. A point outside horopter moves inward.
3. The shift depends on the ratio of the depth of the point with horopter.

Shift in called disparity:

\[ d = x - x' \]

\[ z = \frac{b}{d} \]

Note that this is a relative depth.
Let us assume that same camera moves along the principal axis.

\[ c_1 = k [ I | 0 ] \]

\[ c_2 = k [ I | t_2 ] \quad \text{where} \quad t_2 = \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \]

Let \( k = \begin{pmatrix} k_1 & k_2 & k_3 \\ 0 & k_4 & k_5 \\ 0 & 0 & k_6 \end{pmatrix} \)

\[ p = c_1 \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \]

\[ = \begin{pmatrix} k_1 x + k_2 y + k_3 t \\ k_4 y + k_5 t \\ k_6 z \end{pmatrix} \]

\[ p' = c_2 \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \]

\[ = \begin{pmatrix} k_1 x + k_2 y + k_3 t + k_3 t \\ k_4 y + k_5 t + k_5 t \\ k_6 z + k_6 t \end{pmatrix} \]
Let \( p = \left( \begin{array}{c} x \\ y \\ 1 \end{array} \right) \) and \( p' = \left( \begin{array}{c} x' \\ y' \\ 1 \end{array} \right) \).

\[
\begin{align*}
\therefore x &= \frac{k_1}{k_6} \cdot \frac{x}{k_2} + \frac{k_2}{k_6} \cdot \frac{y}{k_3} + \frac{k_3}{k_4} \\
y &= \frac{k_4}{k_6} \cdot \frac{y}{k_5} + \frac{k_5}{k_6}
\end{align*}
\]

Similarly, you can find \((x', y')\).

Then you will find this eqn.

\[
\begin{align*}
\mathbf{z} - x' &= \frac{t}{k_6} \left( x' - \frac{k_3}{k_6} \right) \\
y - y' &= \frac{t}{k_6} \left( y' - \frac{k_5}{k_6} \right)
\end{align*}
\]

Now note if \( z = \infty \) then

\( p = p' \). This point is called

focus of expansion & it is the

\( e_1 = e_2 \).
Second, if $z$ decreases, the movement of the 3D pt. in the camera increases. i.e., more displacement for closer pts. Than further pt. Also, with the increase in $t$, the displacement is more.

Optical Flow

Note $(x', y')$ is related to $(x, y)$ by a linear equation. Hence all corresponding pts will fall on a line.
So, think about the opposite thing, if you have been able to figure out these lines and given the motif, just by analysing shifting of same pts, you can recover depth. Structure from motif.