Image Stitching

Image Stitching

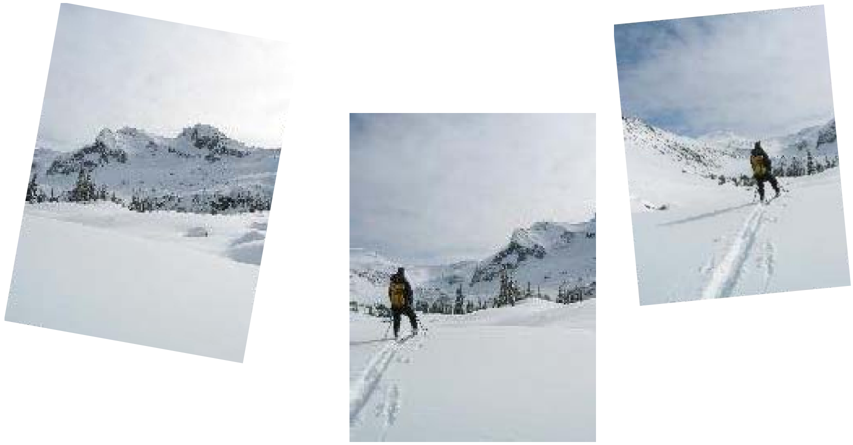
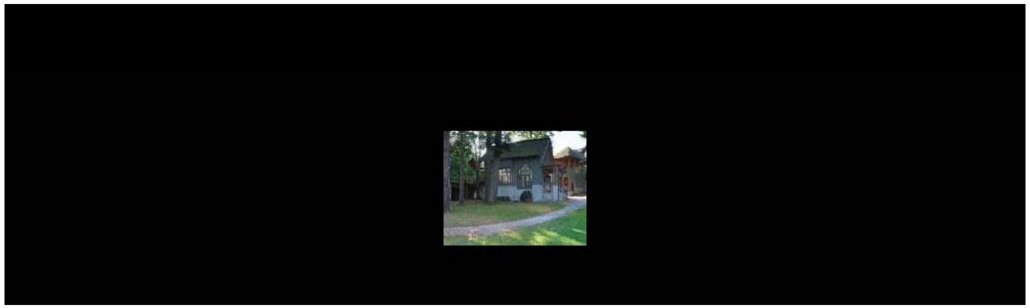


Image Stitching

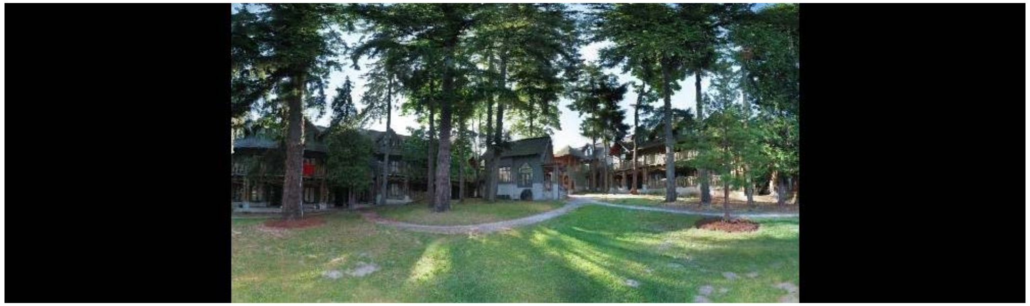
Field of view

- Are you getting the whole picture?
  - Compact Camera FOV = 50 × 35° (horizontal × vertical)



Field of view

- Are you getting the whole picture?
  - Compact Camera FOV = 50 × 35° (horizontal × vertical)
  - Human FOV = 200 × 135°



Field of view

- Are you getting the whole picture?
  - Compact Camera FOV = 50 × 35°
  - Human FOV = 200 × 135°
  - Panoramic Mosaic = 360 × 180°

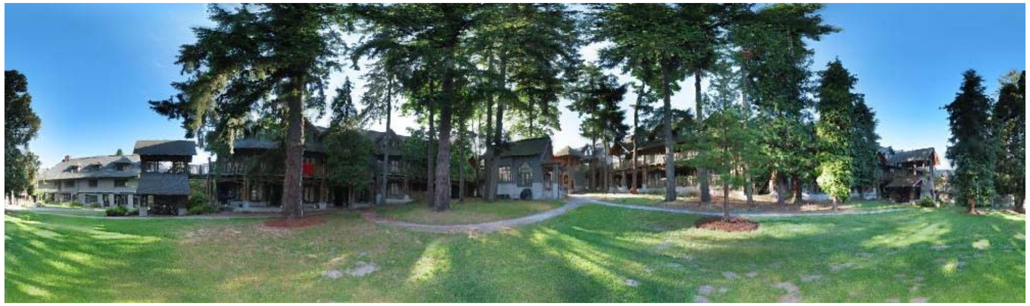


Image Alignment

- 1D Rotations (θ)
  - Ordering ⇒ matching images

Image Alignment

- 1D Rotations (θ)
  - Ordering ⇒ matching images

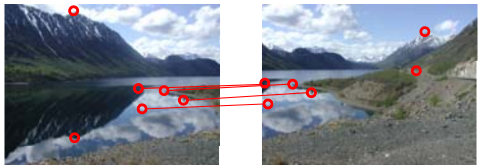
Image Alignment

- 2D Rotations (θ, ϕ)
  - Ordering $\Rightarrow$ matching images

Image Alignment

- 2D Rotations (θ, ϕ)
  - Ordering $\Rightarrow$ matching images

Image Stitching Algorithm Overview



1) **Keypoint Detection**
2) **Keypoint Matching**
3) **Transformation Estimation**
4) **Image Warping**
5) **Image Blending**

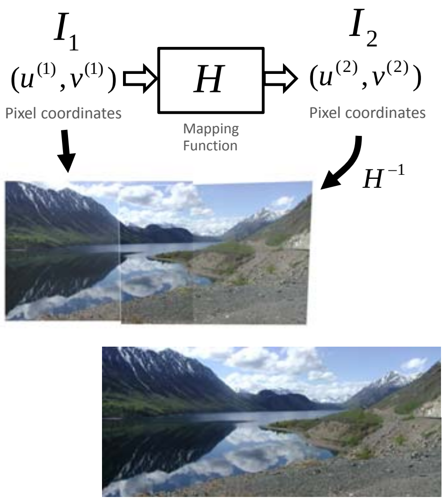
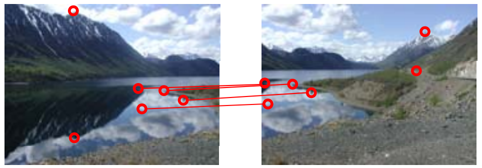


Image Stitching Algorithm Overview



Matched Keypoints

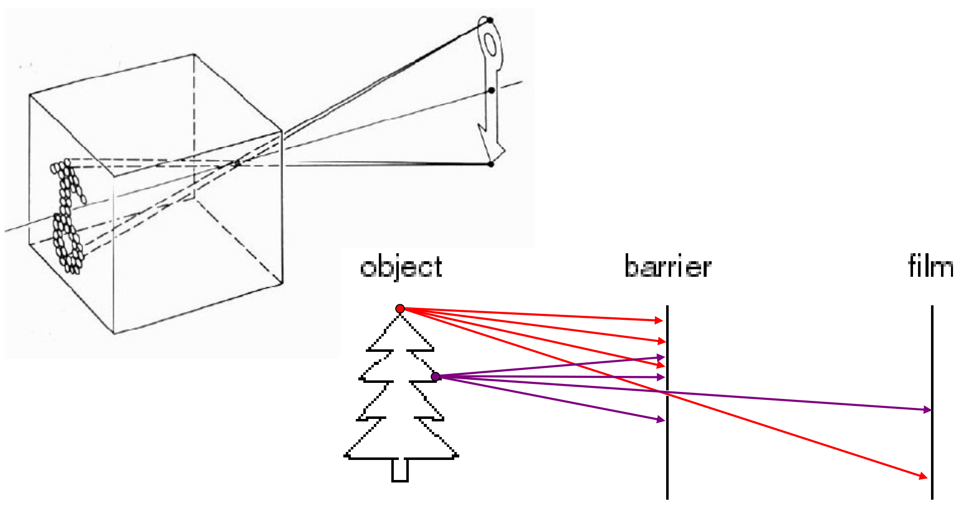
I_1 | I_2

$$(u_1^{(1)}, v_1^{(1)}) \Leftrightarrow (u_1^{(2)}, v_1^{(2)})$$
$$(u_2^{(1)}, v_2^{(1)}) \Leftrightarrow (u_2^{(2)}, v_2^{(2)})$$
$$\vdots$$
$$(u_K^{(1)}, v_K^{(1)}) \Leftrightarrow (u_K^{(2)}, v_K^{(2)})$$

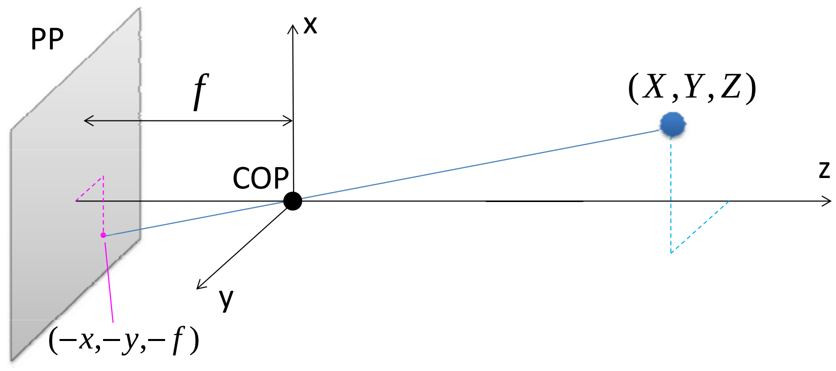
Want to find this transformation function !!!!

$$(u^{(2)}, v^{(2)}) = H\{(u^{(1)}, v^{(1)})\}$$
$$\Downarrow$$
$$(u^{(1)}, v^{(1)}) = H^{-1}\{(u^{(2)}, v^{(2)})\}$$

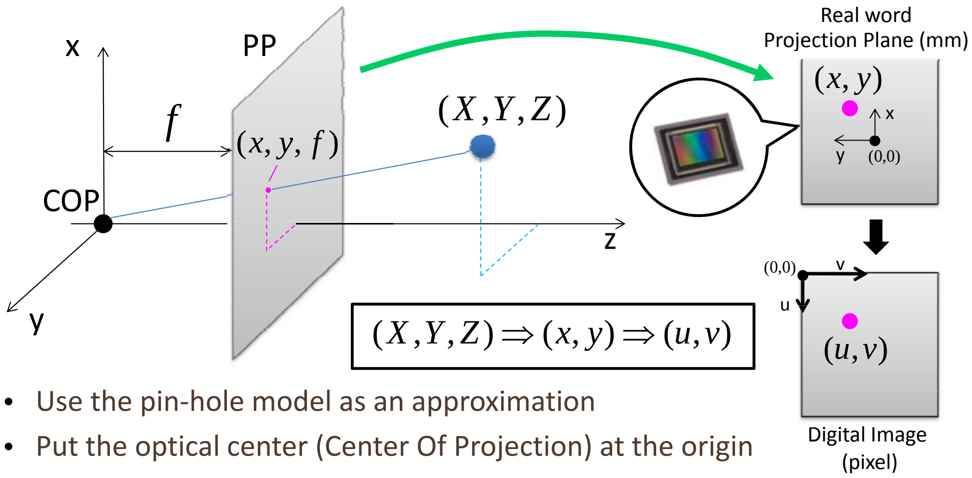
Pin-hole Camera



Pin-hole Camera



Modeling Projection

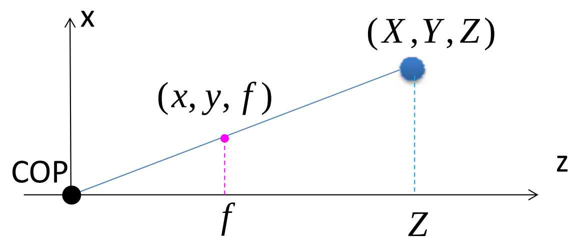


$$\boxed{(X,Y,Z) \Rightarrow (x,y) \Rightarrow (u,v)}$$

- Use the pin-hole model as an approximation
- Put the optical center (Center Of Projection) at the origin
- The camera looks up the positive z axis
- Put the image plane (Projection Plane) in front of the COP

Modeling Projection

$$\boxed{(X,Y,Z) \Rightarrow (x,y)}$$



Similar Triangles

$$\frac{x}{f} = \frac{X}{Z}$$

$$\frac{y}{f} = \frac{Y}{Z}$$

$$(x,y) = \left(\frac{fX}{Z}, \frac{fY}{Z}\right)$$

Non-linear Transformation

Homogenous Coordinates

Add one more coordinate to the Cartesian coordinate:

$$(x, y)_C \Rightarrow (x, y, 1)_H \qquad (X, Y, Z)_C \Rightarrow (X, Y, Z, 1)_H$$

Homogeneous Image Coordinates (2D)

Homogeneous Scene Coordinates (3D)

Converting from homogeneous coordinates to Cartesian coordinate

$$(\tilde{x}, \tilde{y}, \tilde{w})_H \Rightarrow \left(\frac{\tilde{x}}{\tilde{w}}, \frac{\tilde{y}}{\tilde{w}}\right)_C \qquad (\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{w})_H \Rightarrow \left(\frac{\tilde{X}}{\tilde{w}}, \frac{\tilde{Y}}{\tilde{w}}, \frac{\tilde{Z}}{\tilde{w}}\right)_C$$

Projection in Homogenous Coordinates

Cartesian Coordinates

Non-linear Transformation

$$(X, Y, Z)_C \Rightarrow (x, y)_C$$

$$(x, y)_C = \left(\frac{fX}{Z}, \frac{fY}{Z}\right)_C$$

Homogenous Coordinates

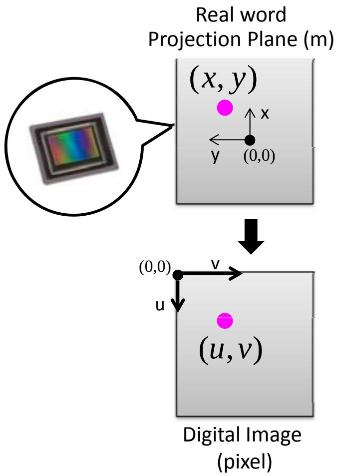
Linear Transformation

$$(X, Y, Z, 1)_H \Rightarrow (\tilde{x}, \tilde{y}, \tilde{w})_H$$

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{w} \end{bmatrix}_H = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_H$$

$$(\tilde{x}, \tilde{y}, \tilde{w})_H = (fX, fY, Z)_H$$

Transform to Digital Image Coordinate System



Apply scaling and translation of origin point (s_u, s_v) (u_0, v_0)

Cartesian Coordinates

$$(x, y)_C \Rightarrow (u, v)_C$$

$$(u, v)_C = (s_u x + u_0, s_v y + v_0)_C$$

Transform to Digital Image Coordinate System

Homogenous Coordinates

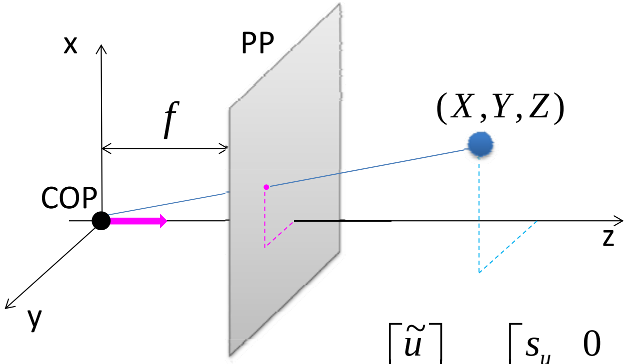
$$(\tilde{x}, \tilde{y}, \tilde{w})_H \Rightarrow (\tilde{u}, \tilde{v}, \tilde{t})_H$$

Scaling Translation

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{t} \end{bmatrix}_H = \begin{bmatrix} s_u & 0 & u_0 \\ 0 & s_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{w} \end{bmatrix}_H \Rightarrow \begin{aligned} \tilde{u} &= s_u \tilde{x} + u_0 \tilde{w} \\ \tilde{v} &= s_v \tilde{y} + v_0 \tilde{w} \\ \tilde{t} &= \tilde{w} \end{aligned}$$

$$(u, v)_C \Rightarrow \left(\frac{\tilde{u}}{\tilde{t}}, \frac{\tilde{v}}{\tilde{t}}\right)_C = \left(\frac{s_u \tilde{x}}{\tilde{w}} + u_0, \frac{s_v \tilde{y}}{\tilde{w}} + v_0\right)_C = (s_u x + u_0, s_v y + v_0)_C$$

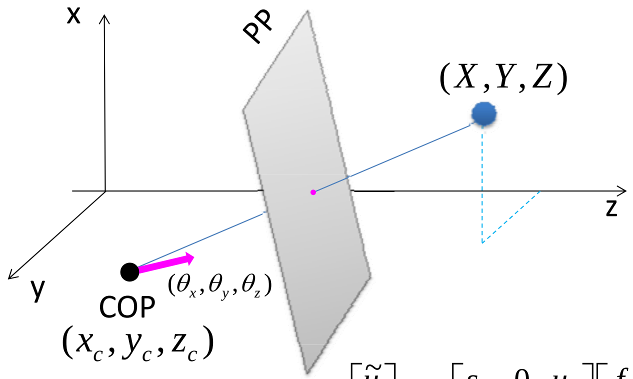
Transform to Digital Image Coordinate System



Camera Position
= (0,0,0)
Camera Direction
= z-axis

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix}_H = \begin{bmatrix} s_u & 0 & u_0\\ 0 & s_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}_H$$

Transform to Digital Image Coordinate System



Camera Position
= (x_c,y_c,z_c)
Camera Direction
= (θ_x,θ_y,θ_z)

Camera Rotation Matrix

Camera Translation Vector

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix}_H = \begin{bmatrix} s_u & 0 & u_0\\ 0 & s_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}\boxed{\mathbf{R}_{3\times3}} & \boxed{\mathbf{t}_{3\times1}}\\ \mathbf{0}_{1\times3} & 1\end{bmatrix}^{-1}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}_H$$

Camera Model

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix}_H = \underbrace{\begin{bmatrix} s_u & 0 & u_0\\ 0 & s_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}}_{\text{Intrinsic}}\underbrace{\begin{bmatrix}\mathbf{R}_{3\times3} & \mathbf{t}_{3\times1}\\ \mathbf{0}_{1\times3} & 1\end{bmatrix}^{-1}}_{\text{Extrinsic}}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}_H$$

$$\mathbf{p} = \mathbf{CP}$$

Camera Model

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14}\\ C_{21} & C_{22} & C_{23} & C_{24}\\ C_{31} & C_{32} & C_{33} & C_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

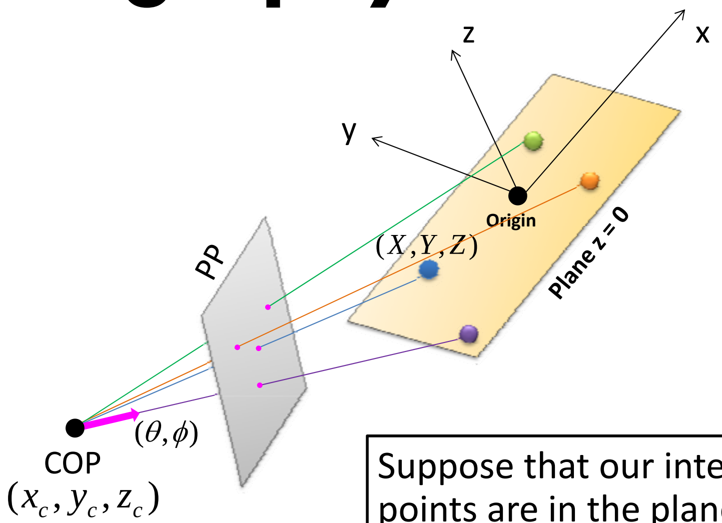
$$(u,v)_C = \left(\frac{\tilde{u}}{\tilde{t}},\frac{\tilde{v}}{\tilde{t}}\right)_C$$

$$\begin{bmatrix}\tilde{u}'\\ \tilde{v}'\\ \tilde{t}'\end{bmatrix} = \alpha\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14}\\ C_{21} & C_{22} & C_{23} & C_{24}\\ C_{31} & C_{32} & C_{33} & C_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

$$(u,v)_C = \left(\frac{\tilde{u}'}{\tilde{t}'},\frac{\tilde{v}'}{\tilde{t}'}\right)_C = \left(\frac{\alpha\tilde{u}}{\alpha\tilde{t}},\frac{\alpha\tilde{v}}{\alpha\tilde{t}}\right)_C = \left(\frac{\tilde{u}}{\tilde{t}},\frac{\tilde{v}}{\tilde{t}}\right)_C$$

- Scale doesn't affect in homogenous camera model
- However $\mathbf{C}$ is non-invertible matrix

Homography



Suppose that our interested points are in the plane z = 0

Homography

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & C_{13} & C_{14}\\ C_{21} & C_{22} & C_{23} & C_{24}\\ C_{31} & C_{32} & C_{33} & C_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

$\Downarrow$

$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & \boxed{C_{13}} & C_{14}\\ C_{21} & C_{22} & \boxed{C_{23}} & C_{24}\\ C_{31} & C_{32} & \boxed{C_{33}} & C_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ \boxed{0}\\ 1\end{bmatrix}$$

Now it is invertible

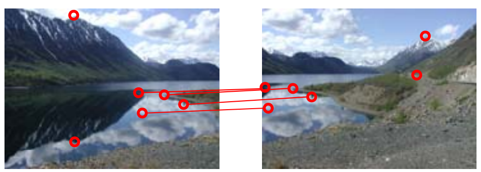
$$\begin{bmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{t}\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & C_{14}\\ C_{21} & C_{22} & C_{24}\\ C_{31} & C_{32} & C_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ 1\end{bmatrix}$$

$\Downarrow$

$$\mathbf{p} = \mathbf{HP}$$

Homography Matrix

Homography Estimation



I_1 I_2

Matched Keypoints

$(u_1^{(1)}, v_1^{(1)}) \Leftrightarrow (u_1^{(2)}, v_1^{(2)})$

$(u_2^{(1)}, v_2^{(1)}) \Leftrightarrow (u_2^{(2)}, v_2^{(2)})$

$\vdots$

$(u_K^{(1)}, v_K^{(1)}) \Leftrightarrow (u_K^{(2)}, v_K^{(2)})$

$$\mathbf{p}_i^{(1)} \Leftrightarrow \mathbf{p}_i^{(2)}$$

$$\mathbf{p}_i^{(1)} = \mathbf{H}^{(1)}\mathbf{P}_i$$

$$\mathbf{p}_i^{(2)} = \mathbf{H}^{(2)}\mathbf{P}_i$$

Same scene point $[X,Y,Z]$

From different camera positions and rotations

$$\mathbf{P}_i = \left(\mathbf{H}^{(1)}\right)^{-1}\mathbf{p}_i^{(1)}$$

$$\mathbf{p}_i^{(2)} = \mathbf{H}^{(2)}\left(\mathbf{H}^{(1)}\right)^{-1}\mathbf{p}_i^{(1)}$$

Homography Estimation

$$\mathbf{p}_i^{(2)} = \mathbf{H}^{(2)}\left(\mathbf{H}^{(1)}\right)^{-1}\mathbf{p}_i^{(1)}$$

$$\mathbf{p}_i^{(2)} = \mathbf{H}_{21}\mathbf{p}_i^{(1)}$$

$$\begin{bmatrix}\tilde{u}^{(2)}\\ \tilde{v}^{(2)}\\ \tilde{t}^{(2)}\end{bmatrix}_H = \begin{bmatrix}h_1 & h_2 & h_3\\ h_4 & h_5 & h_6\\ h_7 & h_8 & h_9\end{bmatrix}\begin{bmatrix}\tilde{u}^{(1)}\\ \tilde{v}^{(1)}\\ 1\end{bmatrix}_H$$

Homography Estimation

$$\begin{bmatrix} \tilde{u}^{(2)} \\ \tilde{v}^{(2)} \\ \tilde{t}^{(2)} \end{bmatrix}_H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} \tilde{u}^{(1)} \\ \tilde{v}^{(1)} \\ 1 \end{bmatrix}_H$$

- Scale doesn't affect in homogenous camera model
- We can set one element to 1
- Actually there are 8 unknown parameters in camera model

Homography Estimation

$$\begin{bmatrix} \tilde{u}^{(2)} \\ \tilde{v}^{(2)} \\ \tilde{t}^{(2)} \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} \tilde{u}^{(1)} \\ \tilde{v}^{(1)} \\ 1 \end{bmatrix}$$

$$(u^{(1)}, v^{(1)})_C = (\tilde{u}^{(1)}, \tilde{v}^{(1)})$$

$$(u^{(2)}, v^{(2)})_C = \left(\frac{\tilde{u}^{(2)}}{\tilde{t}^{(2)}}, \frac{\tilde{v}^{(2)}}{\tilde{t}^{(2)}}\right)$$

$$u^{(2)} = \frac{\tilde{u}^{(2)}}{\tilde{t}^{(2)}} = \frac{h_1\tilde{u}^{(1)} + h_2\tilde{v}^{(1)} + h_3}{h_7\tilde{u}^{(1)} + h_8\tilde{v}^{(1)} + 1}$$

$$v^{(2)} = \frac{\tilde{v}^{(2)}}{\tilde{t}^{(2)}} = \frac{h_4\tilde{u}^{(1)} + h_5\tilde{v}^{(1)} + h_6}{h_7\tilde{u}^{(1)} + h_8\tilde{v}^{(1)} + 1}$$

Homography Estimation

$$u^{(2)} = \frac{h_1 u^{(1)} + h_2 v^{(1)} + h_3}{h_7 u^{(1)} + h_8 v^{(1)} + 1} \qquad v^{(2)} = \frac{h_4 u^{(1)} + h_5 v^{(1)} + h_6}{h_7 u^{(1)} + h_8 v^{(1)} + 1}$$

$$u^{(2)}(h_7 u^{(1)} + h_8 v^{(1)} + 1) - (h_1 u^{(1)} + h_2 v^{(1)} + h_3) = 0$$

$$v^{(2)}(h_7 u^{(1)} + h_8 v^{(1)} + 1) - (h_4 u^{(1)} + h_5 v^{(1)} + h_6) = 0$$

- Eight unknowns in homography (h_1-h_8)
- Two linear equations can be obtained from a pair of matched keypoints

Homography Estimation

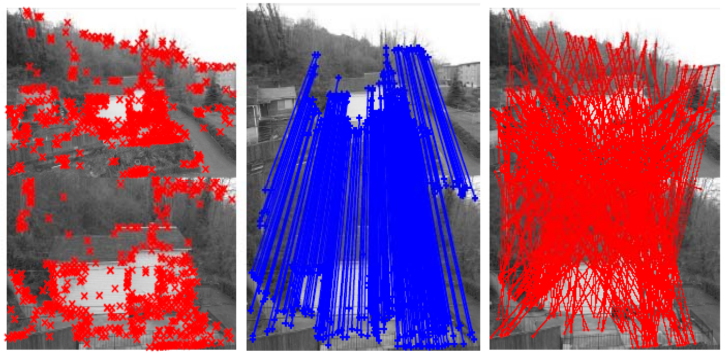
- To estimate homography, we need at least 4 pairs of matched keypoints (8 equations)

$$\left.\begin{array}{c}
-u_1^{(1)}h_1 - v_1^{(1)}h_2 - h_3 + u_1^{(1)}u_1^{(2)}h_7 + v_1^{(1)}u_1^{(2)}h_8 + u_1^{(2)} = 0 \\
-u_1^{(1)}h_4 - v_1^{(1)}h_5 - h_6 + u_1^{(1)}v_1^{(2)}h_7 + v_1^{(1)}v_1^{(2)}h_8 + v_1^{(2)} = 0 \\
-u_2^{(1)}h_1 - v_2^{(1)}h_2 - h_3 + u_2^{(1)}u_2^{(2)}h_7 + v_2^{(1)}u_2^{(2)}h_8 + u_2^{(2)} = 0 \\
-u_2^{(1)}h_4 - v_2^{(1)}h_5 - h_6 + u_2^{(1)}v_2^{(2)}h_7 + v_2^{(1)}v_2^{(2)}h_8 + v_2^{(2)} = 0 \\
\vdots \\
-u_K^{(1)}h_1 - v_K^{(1)}h_2 - h_3 + u_K^{(1)}u_K^{(2)}h_7 + v_K^{(1)}u_K^{(2)}h_8 + u_K^{(2)} = 0 \\
-u_K^{(1)}h_4 - v_K^{(1)}h_5 - h_6 + u_K^{(1)}v_K^{(2)}h_7 + v_K^{(1)}v_K^{(2)}h_8 + v_1^{(2)} = 0
\end{array}\right\}$$

Solve this linear equations then we can obtain $\mathbf{H}$

Outliers

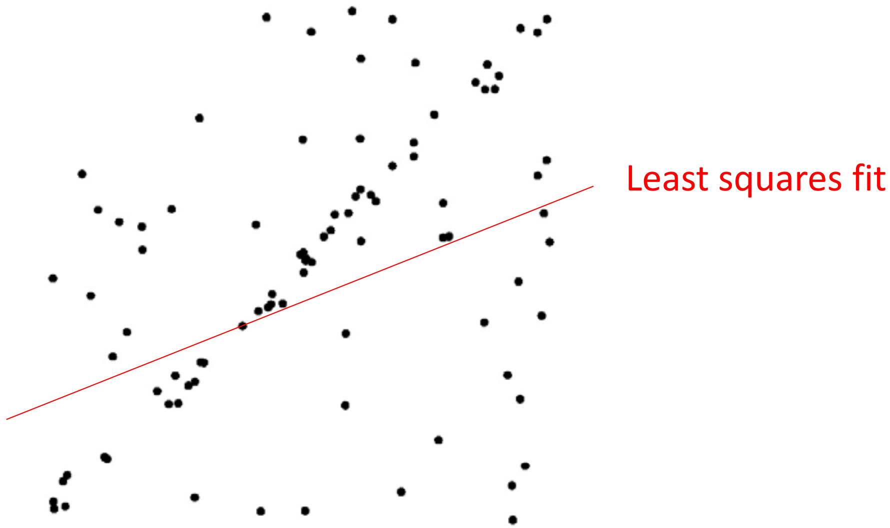
- Assume we have matched points with outliers: How do we compute homography **H**?



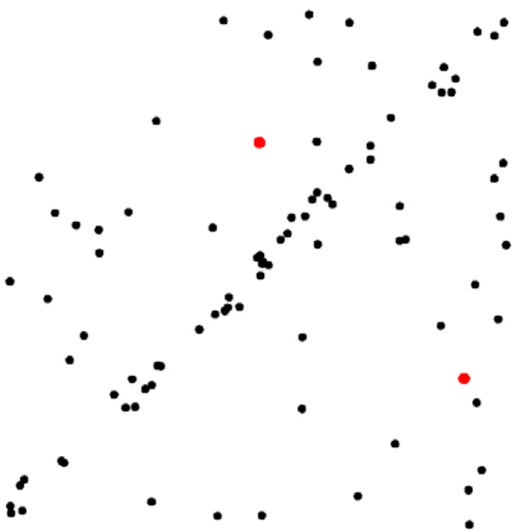
(a) correspondences. (b) identified inliers. (c) identified outliers.

HZ Tutorial '99

RANSAC [RANdom SAmple Concensus]

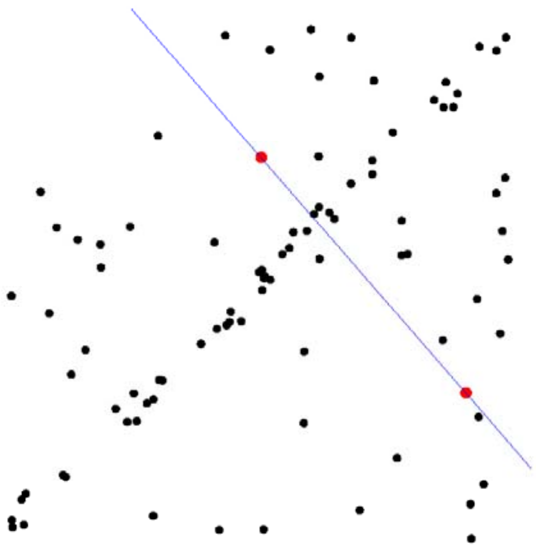


RANSAC [RANdom SAmple Concensus]



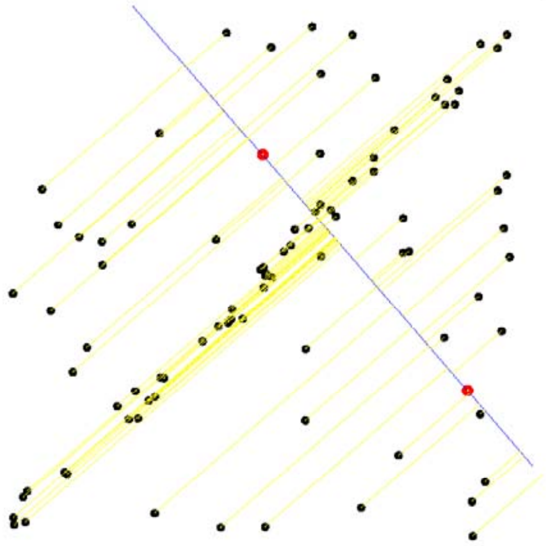
- **Select sample of m points at random**

RANSAC [RANdom SAmple Concensus]



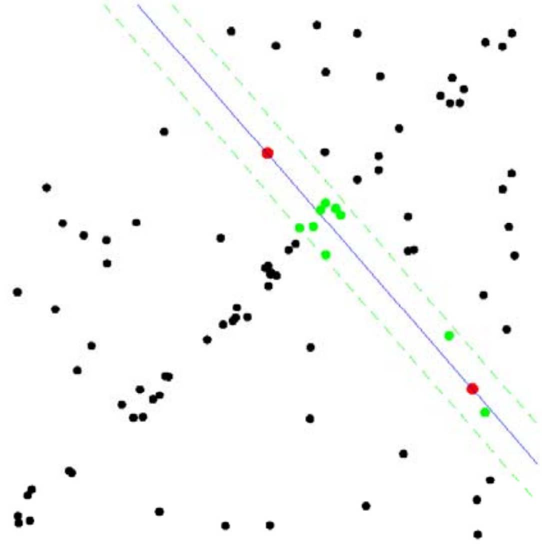
- Select sample of m points at random
- **Calculate model parameters that fit the data in the sample**

RANSAC [RANdom SAmple Concensus]



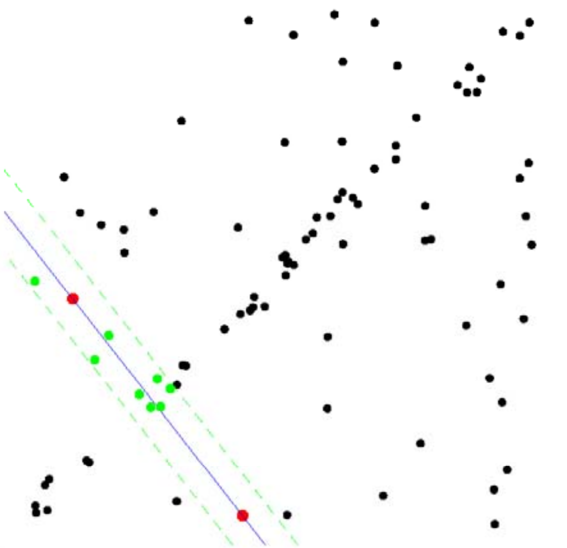
- Select sample of m points at random
- Calculate model parameters that fit the data in the sample
- **Calculate error function for each data point**

RANSAC [RANdom SAmple Concensus]



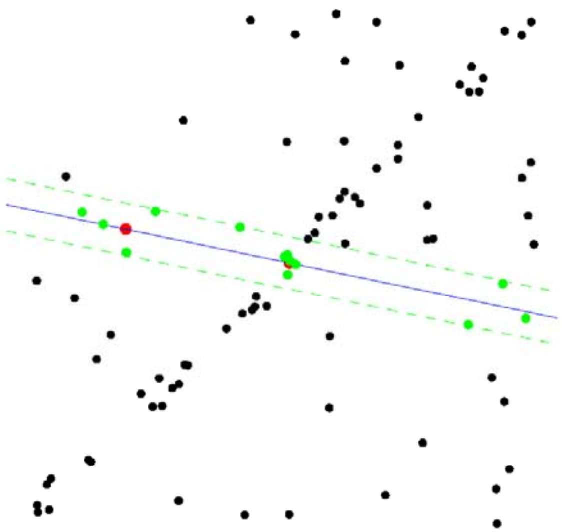
- Select sample of m points at random
- Calculate model parameters that fit the data in the sample
- Calculate error function for each data point
- **Select data that support current hypothesis**

RANSAC [RANdom SAmple Concensus]



- Select sample of m points at random
- Calculate model parameters that fit the data in the sample
- Calculate error function for each data point
- Select data that support current hypothesis
- **Repeat sampling**

RANSAC [RANdom SAmple Consensus]



- Select sample of m points at random
- Calculate model parameters that fit the data in the sample
- Calculate error function for each data point
- Select data that support current hypothesis
- **Repeat sampling**

Automatic Homography Estimation with RANSAC

1. Choose number of samples *N*

For probability p of no outliers:

$$N = \log(1-p)/\log(1-(1-\epsilon)^s)$$

- N, number of samples
- s, size of sample set
- ϵ, proportion of outliers

e.g. for $p = 0.95$

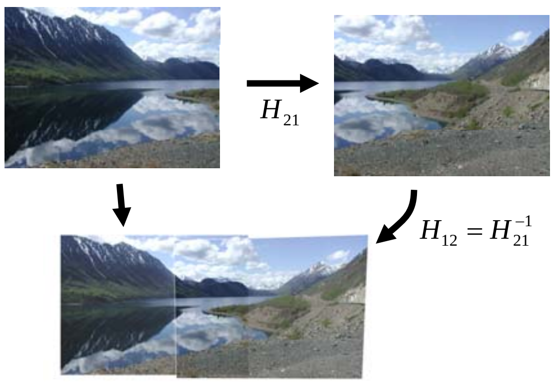
Sample size	Proportion of outliers ϵ						
s	5%	10%	20%	25%	30%	40%	50%
2	2	2	3	4	5	7	11
3	2	3	5	6	8	13	23
4	2	3	6	8	11	22	47
5	3	4	8	12	17	38	95
6	3	4	10	16	24	63	191
7	3	5	13	21	35	106	382
8	3	6	17	29	51	177	766

HZ Tutorial '99

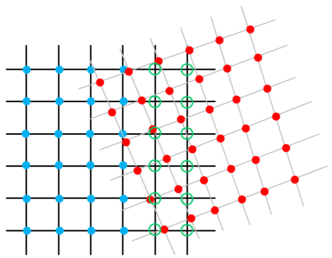
Automatic Homography Estimation with RANSAC

1. Choose number of samples *N*
2. Choose 4 random potential matches
3. Compute **H**
4. Project points from **x** to **x'** for each potentially matching pair:
5. Count points with projected distance < t
   - t ~= 6 σ ; σ is measurement error (1-3 pixels)
6. Repeat steps 2-5 for *N* times
   - Choose **H** with most inliers

Image Warping



$$(u^{(2)}, v^{(2)}) \Rightarrow (u^{(1)}, v^{(1)})$$

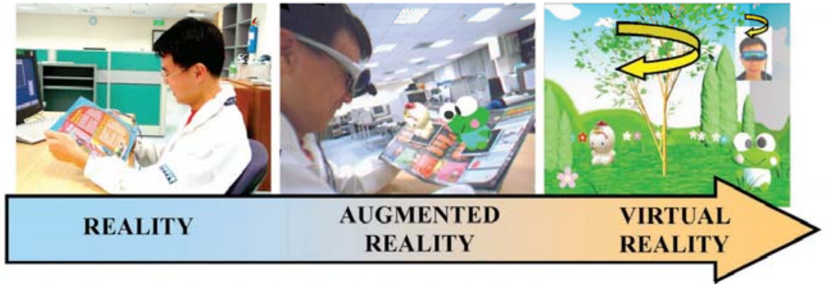


Interpolation

for unknown intensity in the I_1 grid

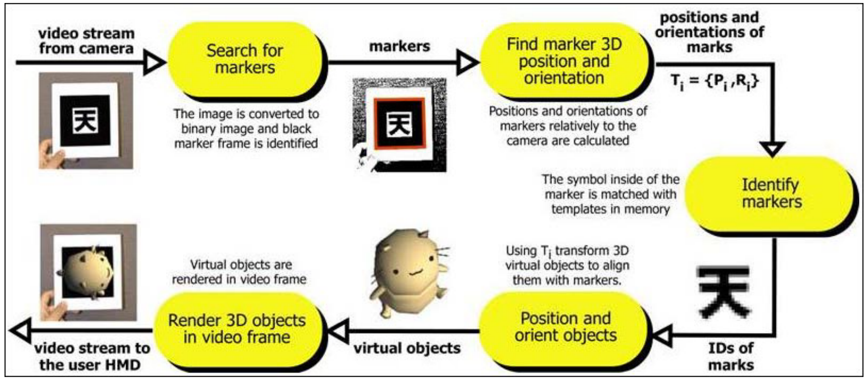
Augmented Reality

AR [Augmented Reality]



http://www0.cs.ucl.ac.uk/staff/ucacsjp/Papers/PrinceXu.pdf

Marker based AR



http://www.hitl.washington.edu/artoolkit/documentation/userarwork.htm

Marker based AR