CITS 4402 Computer Vision

Ajmal Mian

Lecture 09 – Projective Geometry

Overview of this lecture

- Vanishing points and lines
- Projective invariants
- Cross-ratio
- Measuring the height of objects from a single image
- Homography
- Image rectification
- CNN demo and discussion about the Project

Can you work out the geometry of this image?

Notice the pattern of the trees in this 1460 A.D. painting

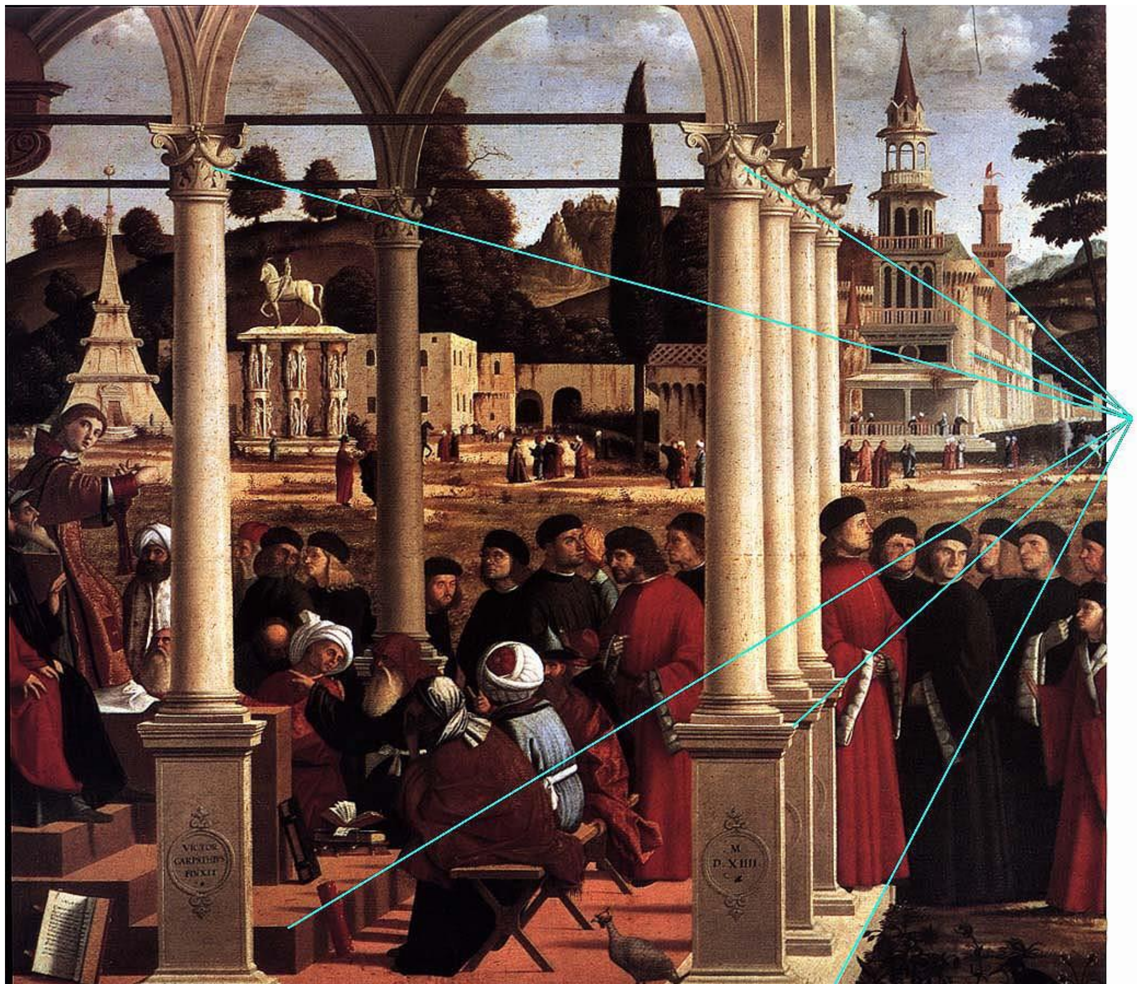
Can you work out the geometry of this image?

Vanishing points

- In perspective image, all parallel lines meet at a single point.
- This point is called the “Vanishing Point”.
- Multiple sets of parallel lines will give multiple vanishing points.

Single vanishing point

Ancient painters were perhaps aware of this concept



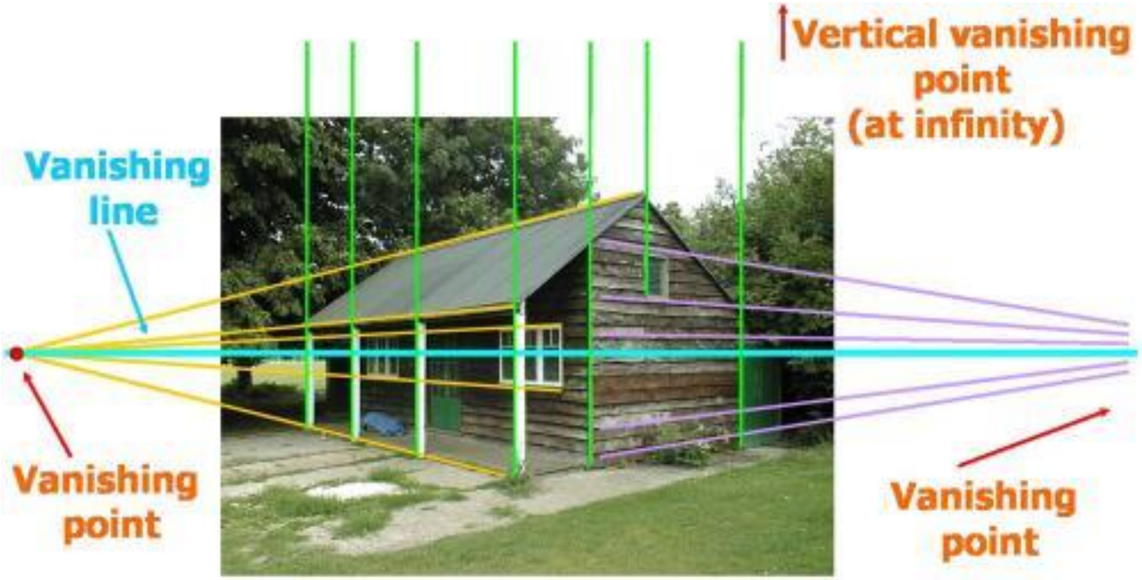
Carpaccio 1514

Two vanishing points

- Can be outside the image

Vanishing points and vanishing line

- Joining two vanishing points give a vanishing line
- Vertical parallel lines give a vertical vanishing point
- Vertical vanishing point is at infinity here



Points at infinity

- What are the image coordinates of the point at infinity on the X-axis?

  A point at infinity on the X-axis is represented by $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$.

  Where will this appear in the image?

$$\begin{bmatrix}su\\sv\\s\end{bmatrix}=\begin{bmatrix}q_{11} & q_{12} & q_{13} & q_{14}\\q_{21} & q_{22} & q_{23} & q_{24}\\q_{31} & q_{32} & q_{33} & q_{34}\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}q_{11}\\q_{21}\\q_{31}\end{bmatrix}$$

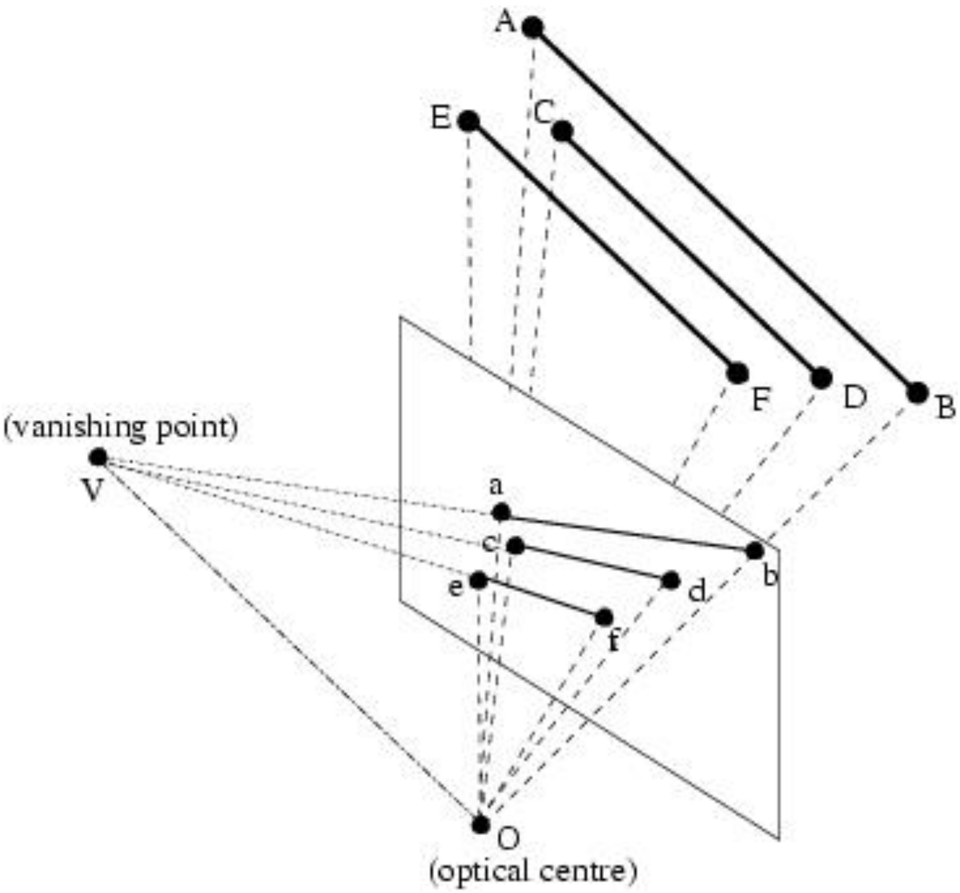
Similarly, points at infinity on the y-axis and z-axis are represented by the 2^{nd} and 3^{rd} columns.

The projection of origin of the world coordinates is given by the last column.

Vanishing points and vanishing lines

- Notice how the parallel lines become non-parallel when projected on the image plane

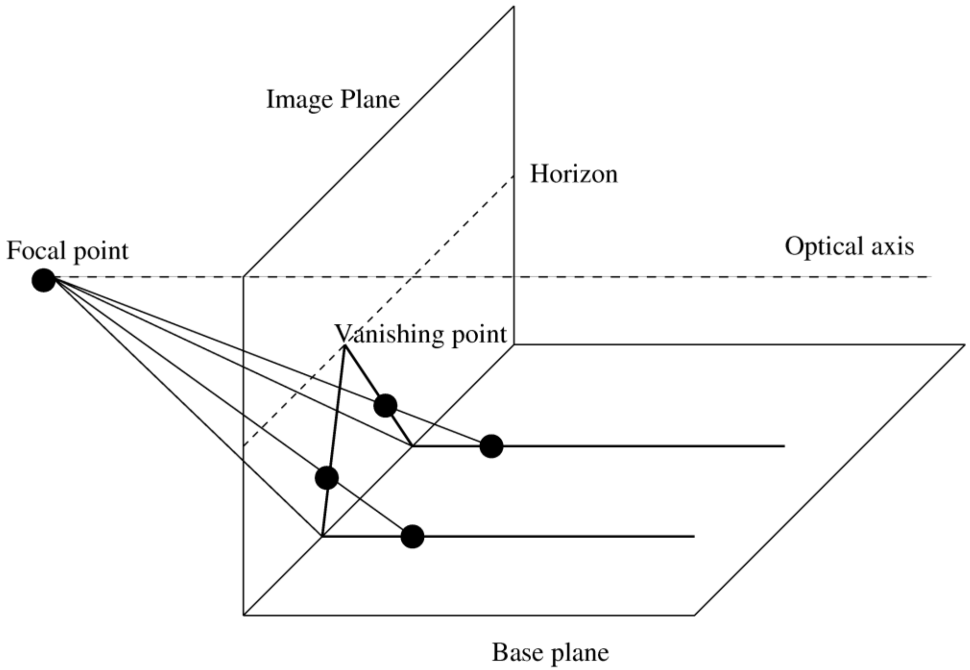
- Vanishing points are often outside the image boundary



Vanishing points and vanishing lines

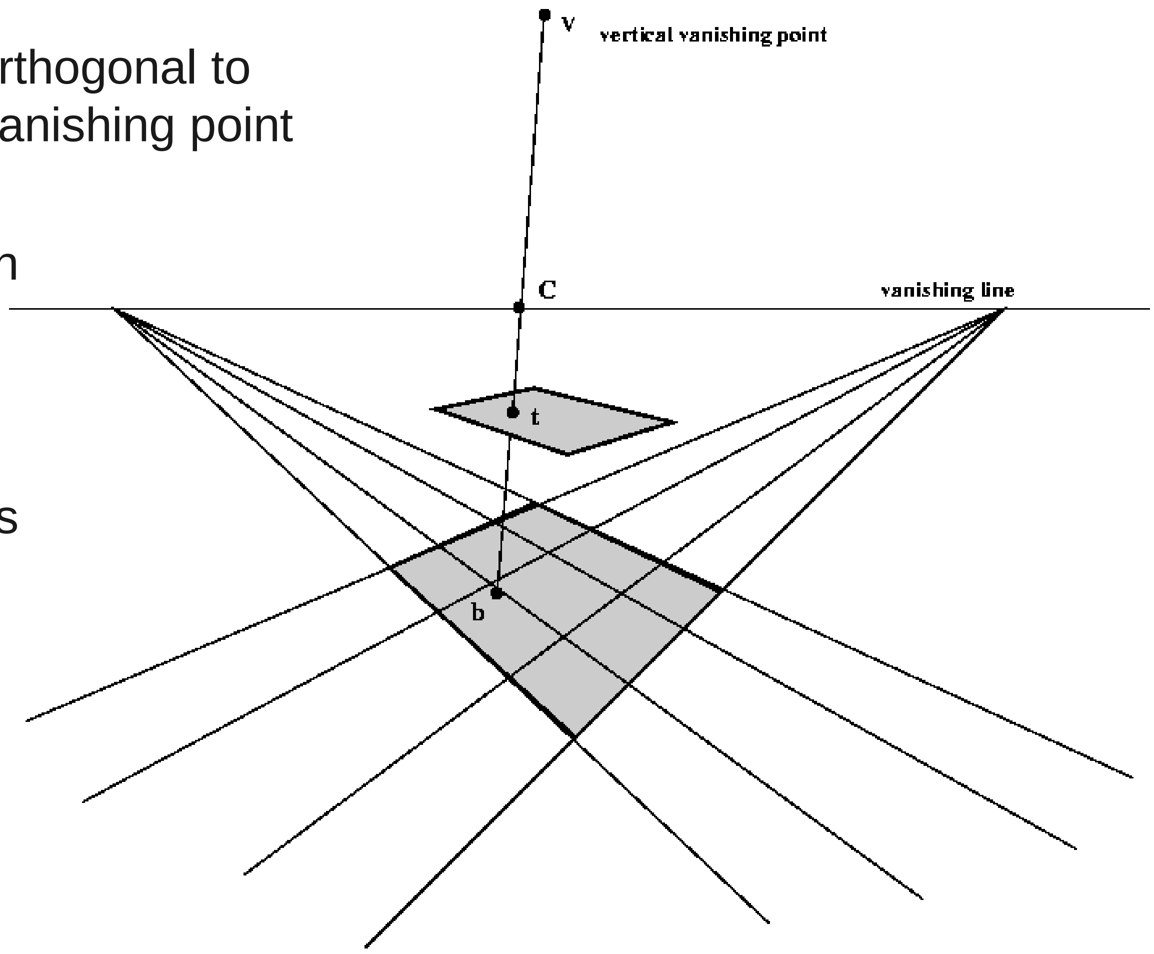
- Notice how the parallel lines become non-parallel when projected on the image plane

- Vanishing points are often outside the image boundary



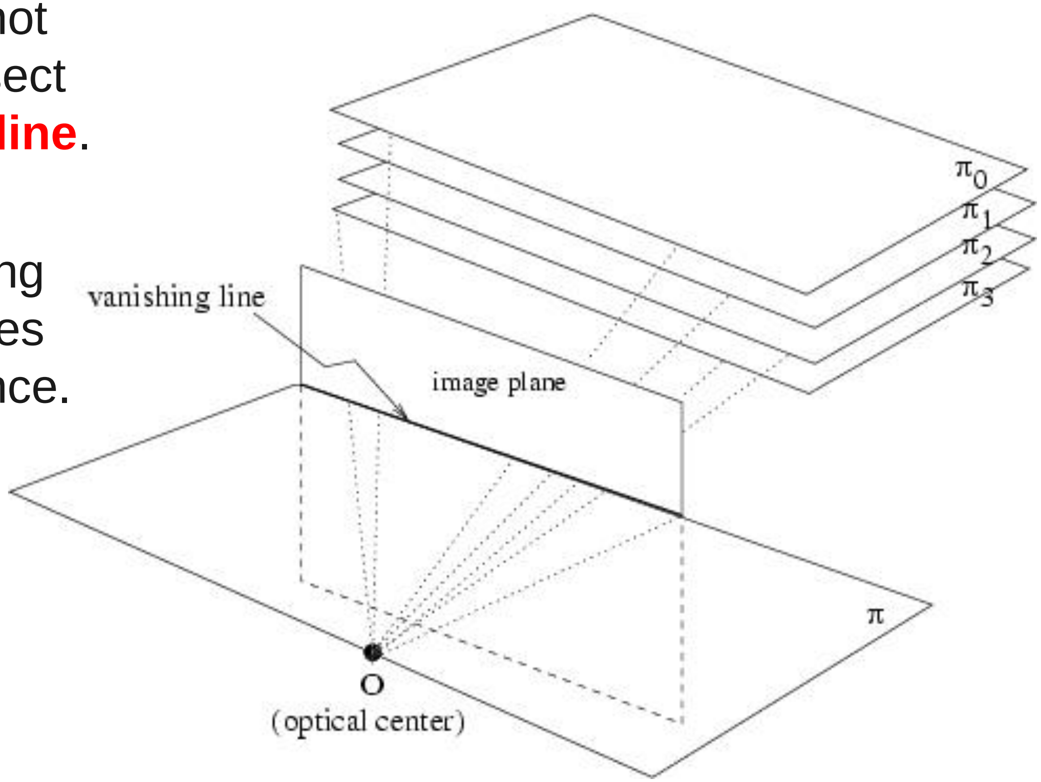
The horizon

- Parallel lines that are not orthogonal to the optical axis meet at a vanishing point
- Two sets of parallel lines on the ground plane will give two vanishing points.
- Joining the vanishing points gives the vanishing line or "**horizon**"



Vanishing lines

- A set of parallel planes that are not parallel to the image plane intersect the image plane at a **vanishing line**.
- The **horizon** is a special vanishing line when the set of parallel planes are parallel to the ground reference.



Anything in the scene that is above the camera will be projected above the horizon in the image.

How to calculate vanishing points and lines

- A point in image is defined by its (u,v) coordinates
- Two points determine a line
- Intersection of two parallel lines will give us a vanishing point
- Two vanishing points will give us the corresponding vanishing line
- Two clicks on each parallel line will solve the problem
- For automatic solution: Edge detection + Hough lines will solve the problem

Points and lines in Homogeneous coordinates

- Points in 2D are represented in homogeneous coordinates as a 3-vector $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$.
- Points and lines are **dual** in 2D, so lines are represented in homogeneous coordinates as 3-vector also: $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$.
- Homogeneous coordinates contain **redundant information** – the representation includes an arbitrary scale!
  e.g., $[2 \quad 3 \quad 1]^T$ and $[4 \quad 6 \quad 2]^T$ represent the same point $[2 \quad 3]^T$ in **inhomogeneous coordinates**.
  e.g., $[4 \quad 7 \quad -3]^T$ and $[2 \quad 3.5 \quad -1.5]^T$ represent the same line – as $4x + 7y - 3 = 0$ and $2x + 3.5y - 1.5 = 0$ denote the same line.
- To determine if a line $\boldsymbol{l} = [a \quad b \quad c]^T$ passes through a point $\boldsymbol{p} = [x \quad y \quad 1]^T$ check if their dot product is 0. That is, $\boldsymbol{l}^T\boldsymbol{p} = 0 \Rightarrow \boldsymbol{l}$ passes through $\boldsymbol{p}$.
  Example: Line $[2 \quad 4 \quad 5]^T$contains the point $[1.5 \quad -2 \quad 1]^T$ since $[2 \quad 4 \quad 5] \cdot [1.5 \quad -2 \quad 1] = 2 * 1.5 - 4 * 2 + 5 * 1 = 0$.

Equation of a line given two points

- The line $\boldsymbol{l}$ that passes through two points $\boldsymbol{p}_1 = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$ and $\boldsymbol{p}_2 = \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$ is obtained via the cross product

$$\boldsymbol{l} = \boldsymbol{p}_1 \times \boldsymbol{p}_2.$$

- **Example:** the line $\boldsymbol{l}$ passing through the points $[0 \quad 2 \quad 1]^T$ and $[3 \quad 0 \quad 1]^T$ is:

$$\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -6 \end{bmatrix}.$$

Thus, the equation of the line is $2x + 3y - 6 = 0$.

Finding the intersection of two lines

- Since lines and points are dual in 2D, we obtain the line coordinates the same way as before. The point of intersection $\boldsymbol{p}$ of two given lines $\boldsymbol{l}_1 = [a_1 \quad b_1 \quad c_1]^T$ and $\boldsymbol{l}_2 = [a_2 \quad b_2 \quad c_2]^T$ is obtained via the cross product

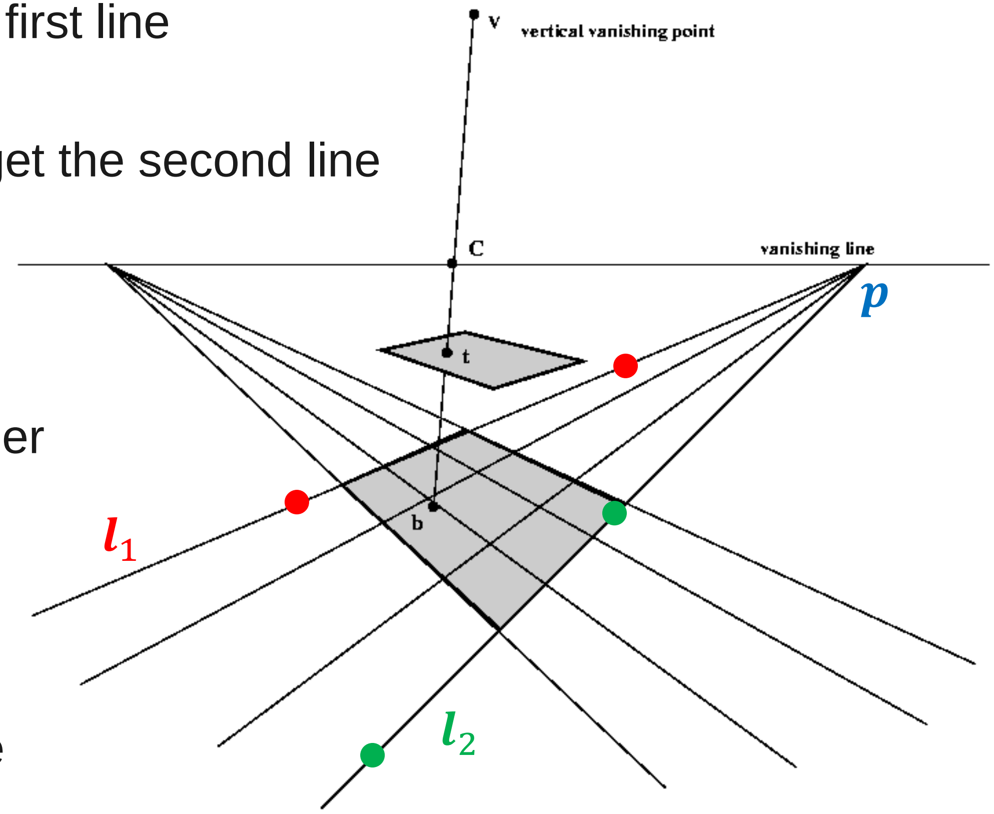
$$\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2.$$

- **Example:** The lines $[4 \quad 6 \quad 2]^T$ and $[2 \quad 0 \quad 1]^T$ intersection at $[6 \quad 0 \quad -12]^T$ which corresponds to the point $\left[-\frac{1}{2} \quad 0\right]^T$ in inhomogeneous coordinates.

- **Exercise:** What are the coordinates of the intersection point of lines $[3 \quad 1 \quad 2]^T$ and $[6 \quad 2 \quad 2]^T$?

Can you find the equation of the vanishing line in this image?

- Click on red points to get the first line
- Click on the green points to get the second line
- Find their intersection
- Repeat the process for another set of parallel lines
- Cross produce of the two vanishing points will give the equation of the vanishing line



What can we find from the vanishing points and lines?

Given the horizon, the vertical vanishing point and one reference height.

The height of any point (from the ground plane) can be computed by specifying the point and its projection on the ground plane in the image.

Criminisi et al 1998.

Human height measurement - example

- If we know the height of one object in the scene
- We can find the height of a person using vanishing points

Representation of various transformations

- **Euclidean transformation:**

$$2D: \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{12} & r_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix} \qquad 3D: \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- **Similarity transformation:**

$$2D: \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{12} & r_{22} & t_2 \\ 0 & 0 & s \end{bmatrix} \qquad 3D: \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & s \end{bmatrix}$$

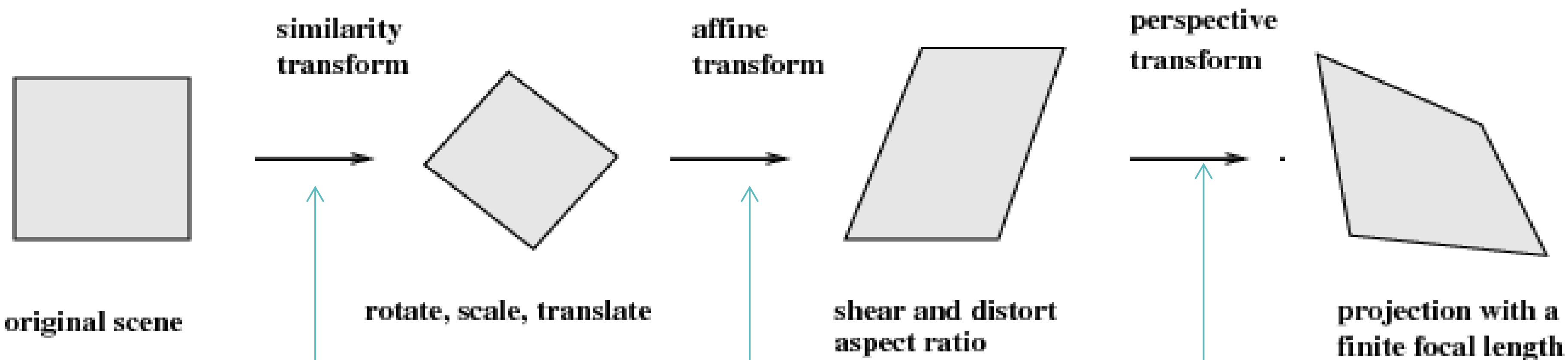
- **Affine transformation:**

$$2D: \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ 0 & 0 & 1 \end{bmatrix} \qquad 3D: \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- **Projective transformation:**

$$2D: \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & 1 \end{bmatrix} \qquad 3D: \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & 1 \end{bmatrix}$$

Geometric invariants



Preserves
- ✓ angles
- ✓ ratios of lengths
- ✓ ratios of areas
- ✓ parallelism
- ✓ concurrency
- ✓ collinearity
- ✓ cross-ratio

Preserves
- ✓ ratios of lengths
- ✓ ratios of areas
- ✓ parallelism
- ✓ concurrency
- ✓ collinearity
- ✓ cross-ratio

Preserves
- ✓ concurrency
- ✓ collinearity
- ✓ cross-ratio

Cross-ratio

- Most important geometric invariant for projective transformation is the **cross-ratio** of 4 points on a line.

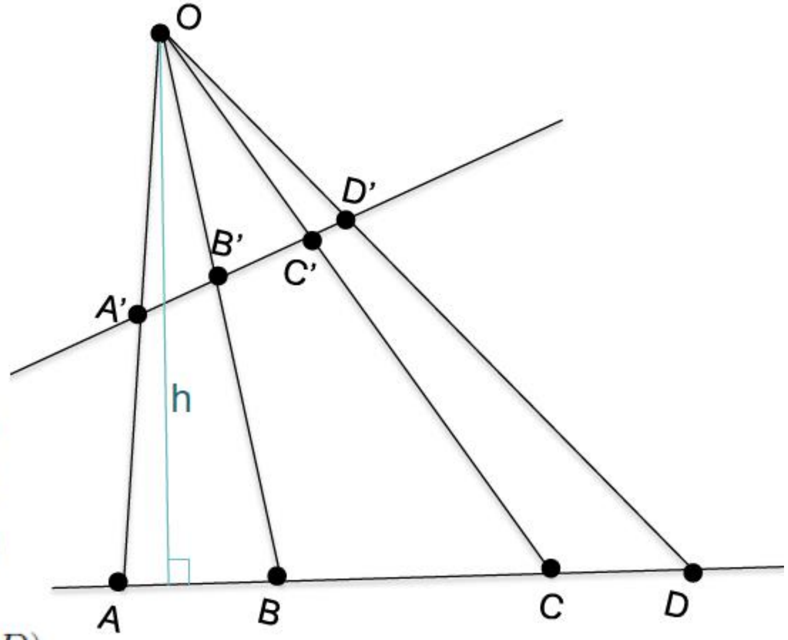
- **The cross-ratio is the ratio of two ratios of lengths**
  - Given 4 points on a line, say A, B, C and D.
  - Take x_1 as the 1st reference point, compute $r_1 = \overrightarrow{AC}/\overrightarrow{AD}$.
  - Take x_2 as the 2nd reference point, compute $r_2 = \overrightarrow{BC}/\overrightarrow{BD}$.
  - Cross-ratio is defined as $r = r_1/r_2$.

- The two reference points and the ratios can also be arbitrary chosen

- e.g. r_1 can be set to $\overrightarrow{CD}/\overrightarrow{CB}$, r_2 can be set to $\overrightarrow{AD}/\overrightarrow{AB}$, and $r = r_2/r_1$.

- There are $4! = 24$ permutations of possible cross-ratio values however, these values are **repeated** leaving only **6 unique cross-ratio possibilities**

Proof of cross-ratio invariance



$$\frac{AC/AD}{BC/BD} = \frac{A'C'/A'D'}{B'C'/B'D'}$$

$$\begin{aligned}
\text{Area}(AOC) &= h/2(AC) = 1/2(OA)(OC)\sin(AOC) \\
\text{Area}(AOD) &= h/2(AD) = 1/2(OA)(OD)\sin(AOD) \\
\text{Area}(BOC) &= h/2(BC) = 1/2(OB)(OC)\sin(BOC) \\
\text{Area}(BOD) &= h/2(BD) = 1/2(OB)(OD)\sin(BOD)
\end{aligned}$$

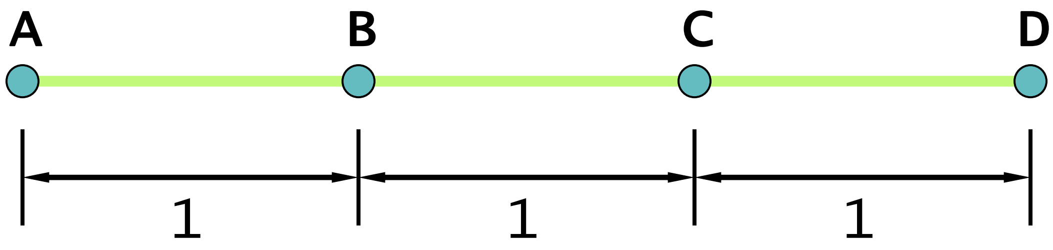
$$\frac{AC/AD}{BC/BD} = \frac{\cancel{(OA)}\cancel{(OC)}\sin(AOC)/\cancel{(OA)}(OD)\sin(AOD)}{(OB)\cancel{(OC)}\sin(BOC)/(OB)(OD)\sin(BOD)}$$

$$\frac{AC/AD}{BC/BD} = \frac{\sin(AOC)/\sin(AOD)}{\sin(BOC)/\sin(BOD)}$$

Cross-ratio depends on the angles and is thus invariant under projection.

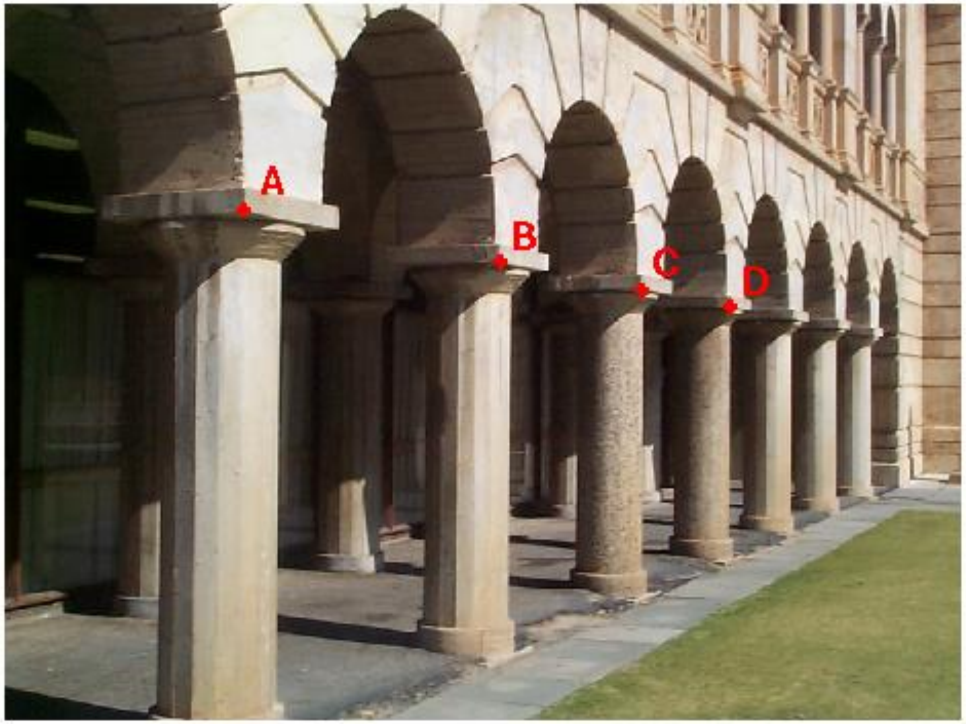
Courant and Robbins, “What is Mathematics?” Oxford University Press 1998

Cross-ratio of equidistant points



$$\text{Cross ratio} = \frac{AC/AD}{BC/BD} = \frac{2/3}{1/2} = 4/3 = 1.333$$

Let us measure it from an image with known equidistant points



- ❑ Cross-ratio can always be calculated from an image.
- ❑ This cross-ratio will be exactly the same as in real world.
- ❑ If one of the measurements is unknown in the real world, it can be calculated using the cross ratio

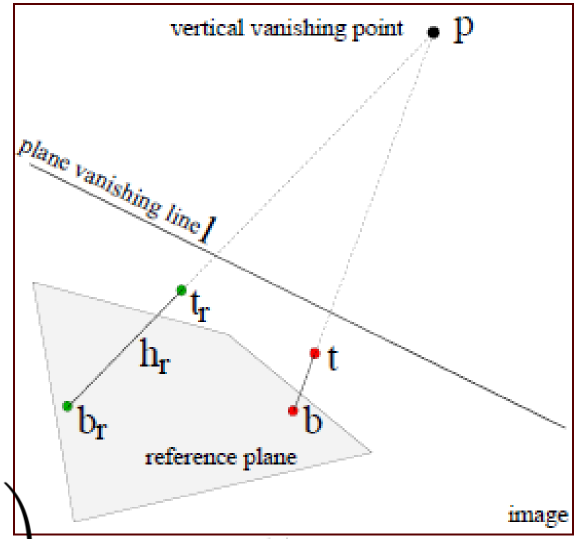
$$\text{Cross ratio} = \frac{AC/AD}{BC/BD} = \frac{488/596}{173/282} = 1.332$$

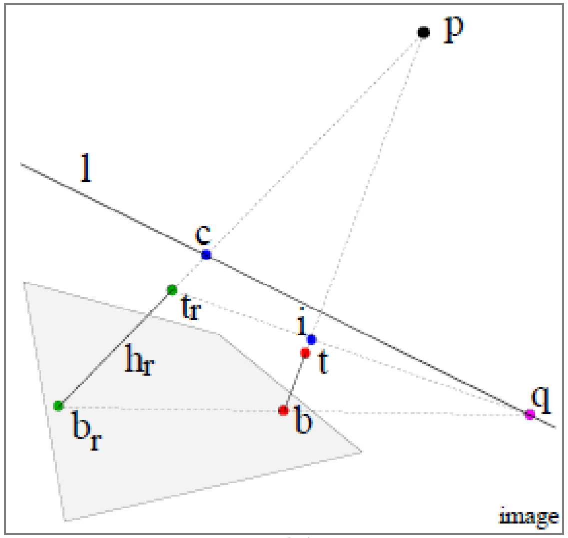
Back to the human height measurement problem

- If we know the height of one object in the scene
- We can find the height of a person using vanishing points

Height measurement with cross-ratio

$$
\begin{aligned}
R_I &= \frac{\vec{bi}/\vec{bp}}{\vec{ti}/\vec{tp}} \\
R_W &= \frac{i/p}{(i-t)/(p-t)} \\
&= \frac{i}{p}\frac{p-t}{i-t} \\
&= \frac{i}{p}\left(\frac{p}{i-t} - \frac{t}{i-t}\right) \\
&= \frac{i}{i-t} - \frac{i}{p}\frac{t}{i-t} \\
&= \frac{i}{i-t}
\end{aligned}
$$





$$
\begin{aligned}
t &= \frac{i(R_W - 1)}{R_W} \\
h &= \frac{h_r(R_I - 1)}{R_I}
\end{aligned}
$$

All distances are signed distances

Algorithm for calculating height

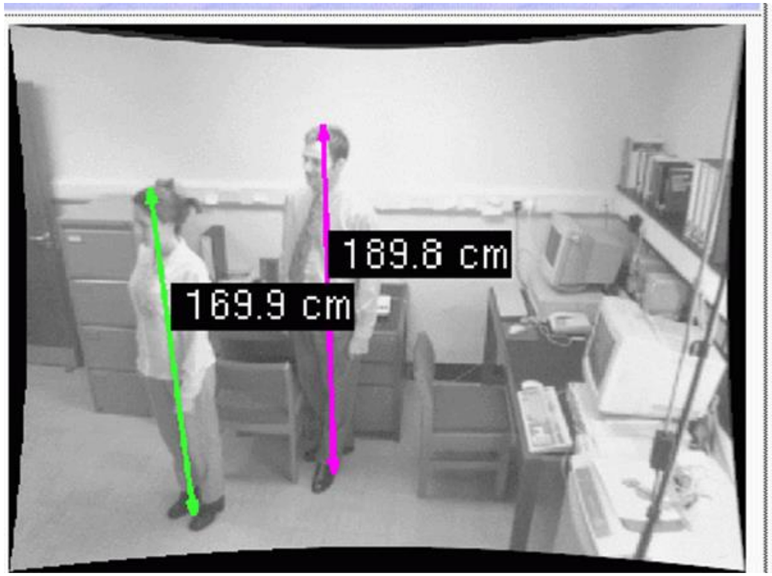
- Calculate vertical vanishing point $\mathbf{v}$
- Calculate vanishing line $\mathbf{l}$ of the reference plane
- Select top $\mathbf{t}_r$ and base points $\mathbf{b}_r$ of the reference object
- Compute metric factor $\alpha = \dfrac{-\|\mathbf{b}_r \times \mathbf{t}_r\|}{H_r(\mathbf{l} \cdot \mathbf{b}_r)\|\mathbf{v} \times \mathbf{t}_r\|}$
- Repeat
  - Select top $\mathbf{t}_x$ and base points $\mathbf{b}_x$ of the object to measure
  - Compute the height H_x of the object

$$H_x = \frac{-\|\mathbf{b}_x \times \mathbf{t}_x\|}{\alpha(\mathbf{l} \cdot \mathbf{b}_x)\|\mathbf{v} \times \mathbf{t}_x\|}$$

A. Criminisi, "Single View Metrology: Algorithms and Applications" 1999.

For precise measurement

- Radial distortion needs to be removed first
- Robust detection of parallel lines
- Vanishing point detection based on multiple parallel lines
- Heights are not always vertical. Ideally, the vertical lines should meet at the vertical vanishing point.



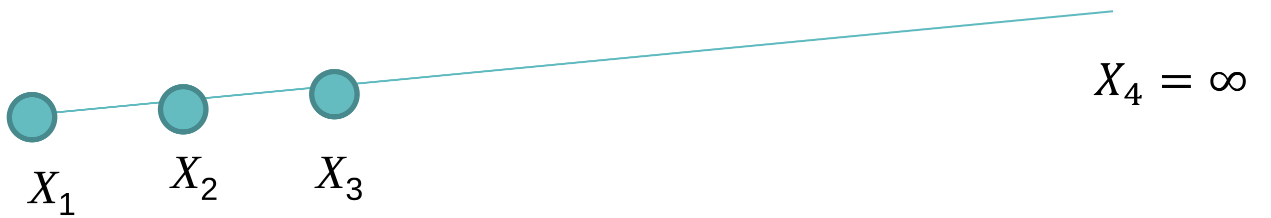
More on cross-ratios

- The 6 possible cross-ratios are: $r, \frac{1}{r}, 1-r, \frac{r-1}{r}, \frac{1}{1-r}, \frac{r}{r-1}$.

- **Example of a special case:** When $r = \frac{1}{2}$. Then the 6 cross-ratios are not unique:

$$\frac{1}{2}, \quad 2, \quad \frac{1}{2}, \quad -1, \quad 2, \quad -1.$$

This corresponds to the situation where one point is the mid-point of the other two points while the 4th point is a point at infinity:



X_2 is the mid-point between X_1 and X_3, i.e., $X_3 - X_2 = X_2 - X_1$.

Homography

- Images of the same planar surface are related by homography
- Homography is used for image rectification (making parallel lines parallel in images)
- Homography is also useful for
  - Image registration
  - Calculation of camera motion (robot navigation, structure from motion)
  - Placing 3D objects in images or videos (augmented reality)
  - Inserting new images such as advertisements in movies

Image rectification

- It is possible to remove perspective distortion of a plane in a scene if
  - We can find the vanishing line of the plane
  - We have two reference measurements of known lengths and angles

- We need to identify 4 points on the plane such that

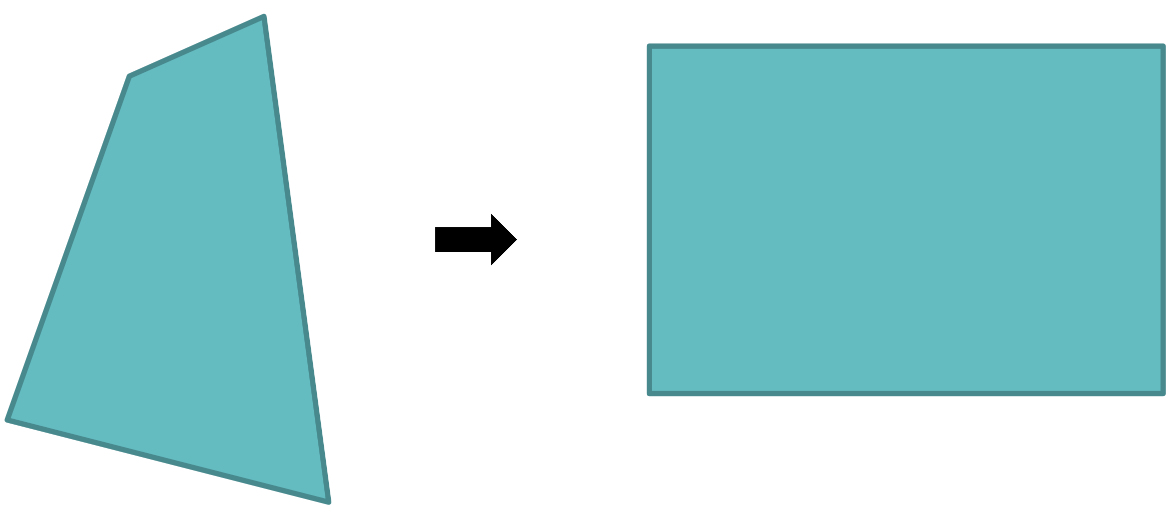
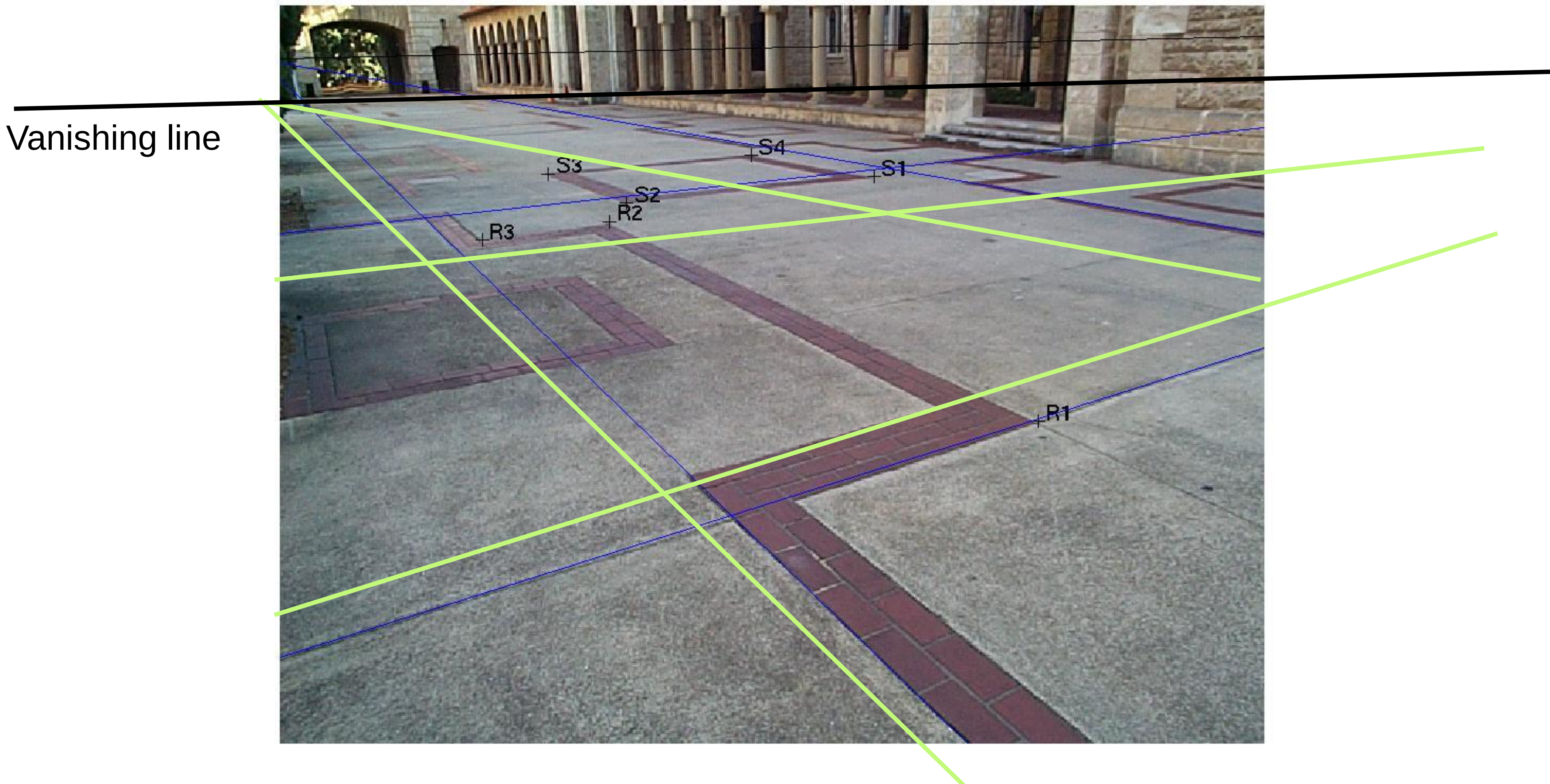


Image rectification example



The points are used to provide constraints on the affine transformation

Rectified image

Cannot include points that are too close to the vanishing line as they are at infinity.

Image rectification

- In general, we need four points for rectifying a plane that has been distorted by a perspective projection.

- Recall the perspective projection equation from Lecture 08

$$s\mathbf{m} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\mathbf{M}$$

where $\mathbf{M} = [X_i, Y_i, Z_i, 1]^T$

- We can expand the above equation $s m_i = \mathbf{K}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t}\end{bmatrix}\begin{bmatrix}X_i \\ Y_i \\ Z_i \\ 1\end{bmatrix}$

Image rectification

- For an xy-plane Z = 0, therefore

$$sm_i = \mathbf{K}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t}\end{bmatrix}\begin{bmatrix}X_i \\ Y_i \\ 0 \\ 1\end{bmatrix}$$

- This simplifies to

$$sm_i = \underbrace{\mathbf{K}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix}}\begin{bmatrix}X_i \\ Y_i \\ 1\end{bmatrix}$$

This 3 x 3 matrix is called
the **Homography Matrix**

Calculating the homography matrix H

- The homography matrix has 9 unknowns and is defined up to an unknown scale

$$\begin{bmatrix} su_i \\ sv_i \\ s \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}$$

- We get

$$u_i = \frac{h_{11}X_i + h_{12}Y_i + h_{13}}{h_{31}X_i + h_{32}Y_i + h_{33}} \qquad v_i = \frac{h_{21}X_i + h_{22}Y_i + h_{23}}{h_{31}X_i + h_{32}Y_i + h_{33}}$$

- Writing it as a system of linear equations

$$\begin{bmatrix} X_i & Y_i & 1 & 0 & 0 & 0 & -u_iX_i & -u_iY_i & -u_i \\ 0 & 0 & 0 & X_i & Y_i & 1 & -v_iX_i & -viYi & -v_i \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ \vdots \\ h_{33} \end{bmatrix} = \mathbf{0}$$

This has for form $Ah = 0$ and the solution is the right singular vector corresponding to the smallest singular value of A i.e. $svd(A) = USV^T$, the last column of V is equal to h.

Homography

- Homography has 8 degrees of freedom, but generally all entries of the 3x3 matrix are treated as unknowns instead of setting one entry to 1

- 4 corresponding are minimum required but generally a large number of correspondences are found automatically (e.g. using the SIFT descriptor)

We need ≥ 4 correspondences, so the matrix A will have ≥ 8 rows

- Some incorrect correspondences are unavoidable

- Thus, robust estimation such as RANSAC is used to compute the homography matrix

Homography – Image mosaicing

- Explained quite well by Thomas Opsahl

http://www.uio.no/studier/emner/matnat/its/UNIK4690/v16/forelesninger/lecture_4_3-estimating-homographies-from-feature-correspondences.pdf

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

- Mosaicing is used to stich aerial images
  - Translate one image to make a bigger image (with black surroundings so we can stick other images there)
  - Find point correspondences between the two images
  - Calculate homography between the two images using a robust method
  - Transform the smaller image to overlay on the bigger one
  - Use smoothing/blending etc to so that individual image edges are not obvious and the lighting is consistent

Affine homography

- **Affine homography** is a more appropriate model if the image region in which the homography is computed is small or the image has been acquired with a large focal length.

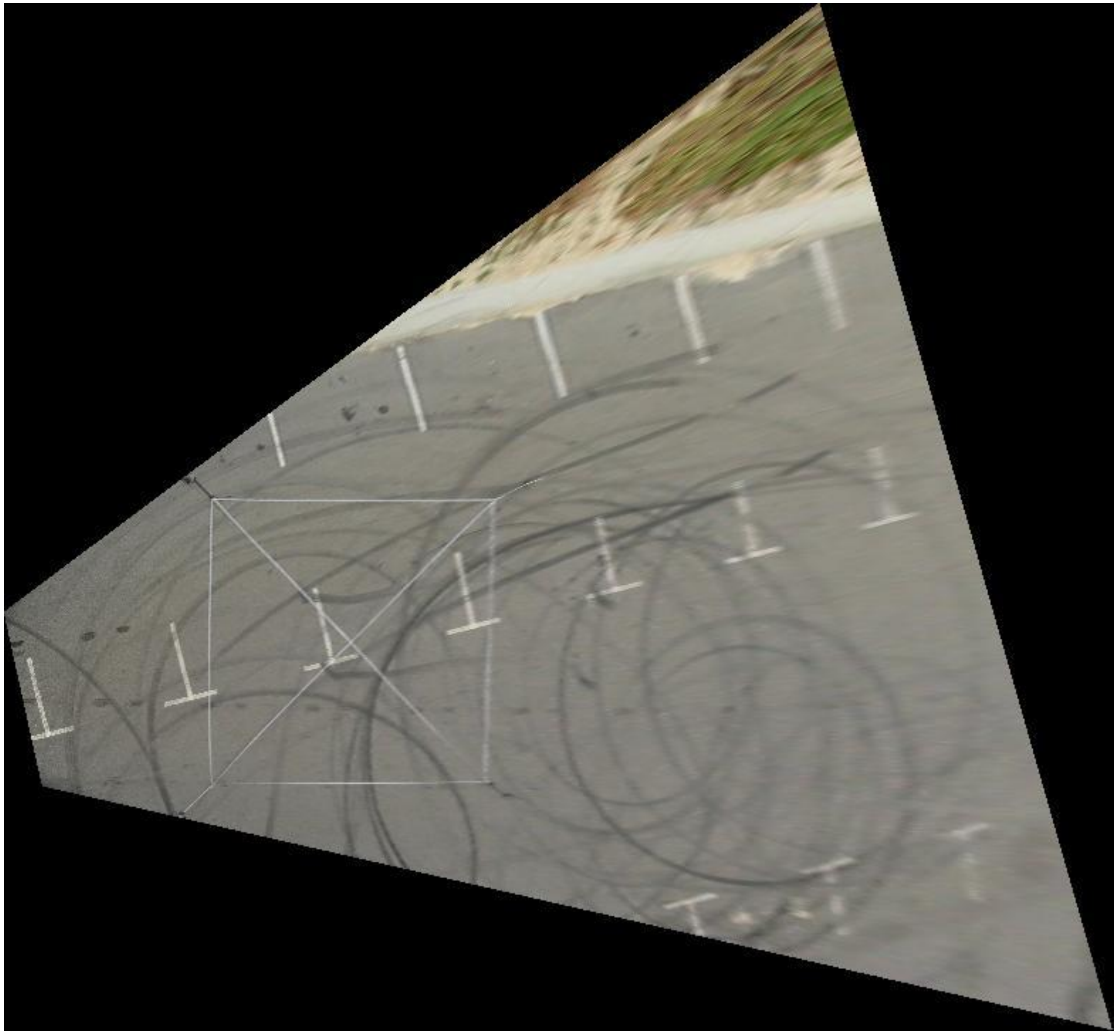
- **Affine homography** is a special type of homography whose last row is fixed to

$$h_{13} = h_{23} = 0, \quad h_{33} = 1$$

- Useful Matlab functions are
  - maketform
  - imtransform
  - estimateGeometricTransform
  - imwarp
  - Also see **homography2d.m** on www.peterkovesi.com

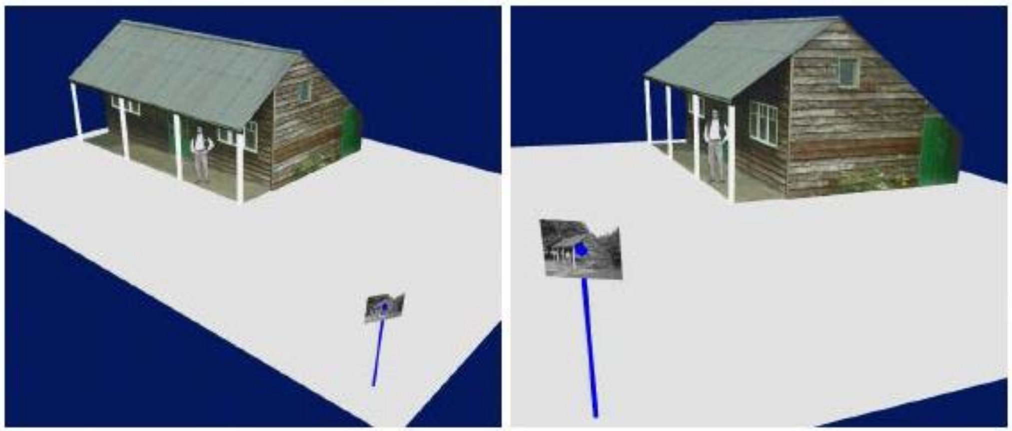
Image rectification example

Image rectification example



Constructing 3D models from single views

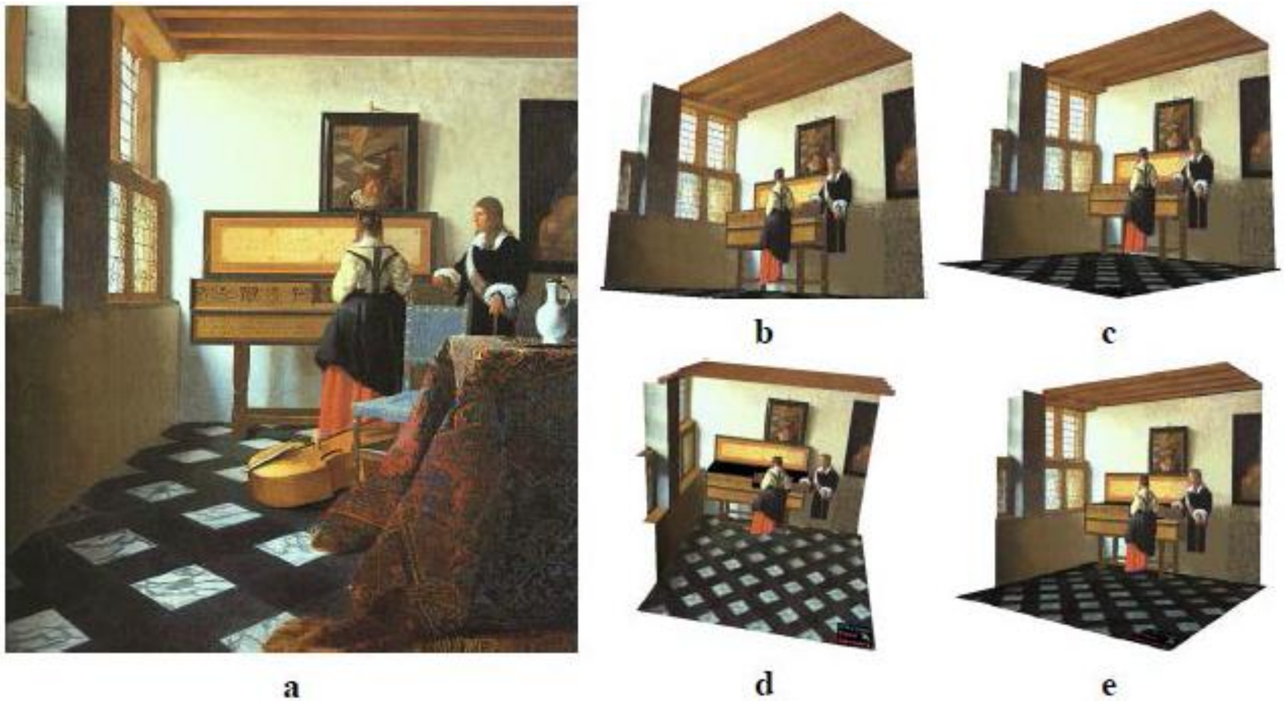
- Extract planes and their relative distances
- Segment scene
- Remove unwanted objects
- Fill occluded area
- The 3D scene mode
  can be rotated



- Note how every segment is approximated with a plane

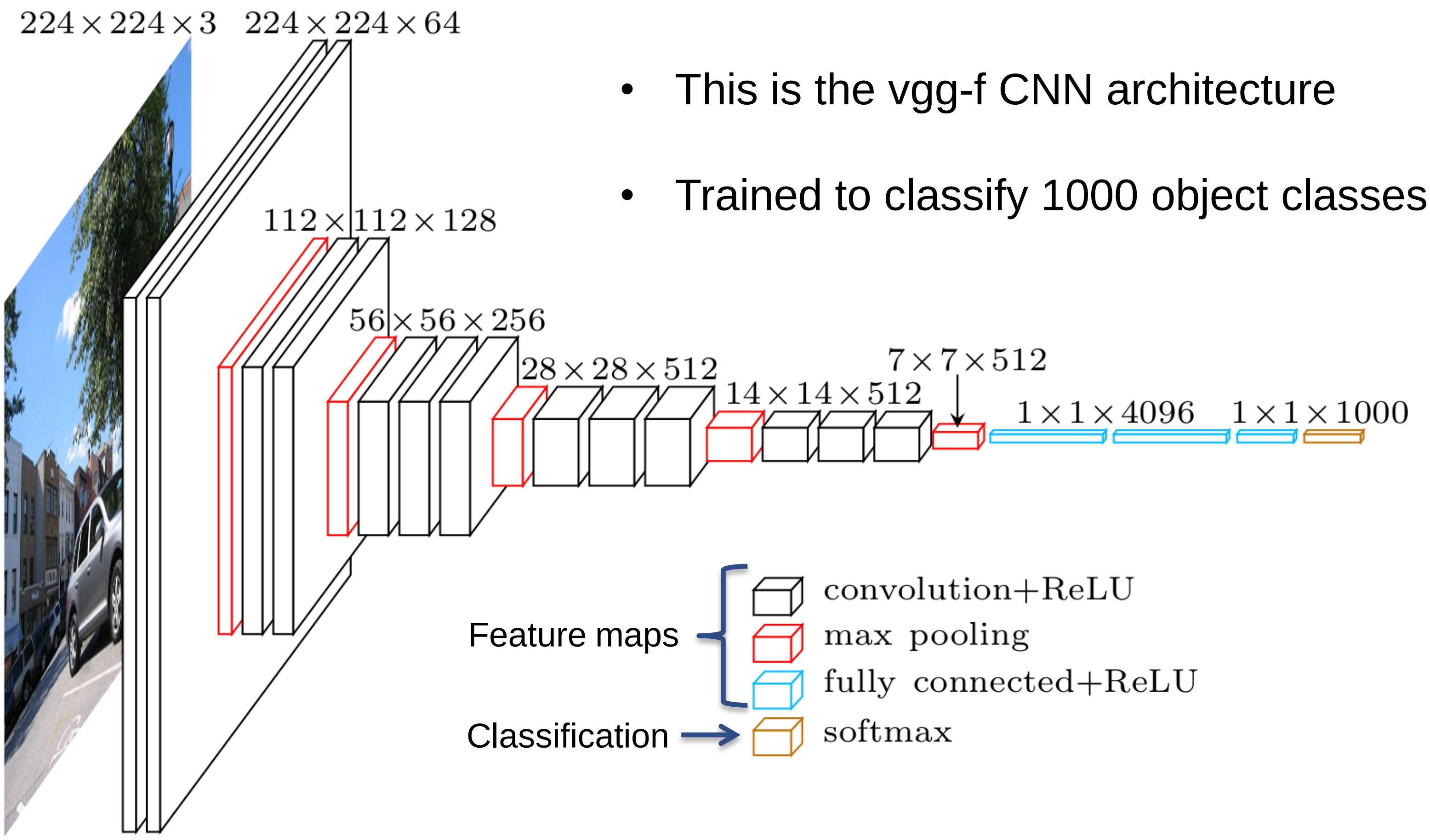
A. Criminisi, “Signle view metrology: Algorithms and Applications”, 1999.

3D reconstruction of historical paintings



A. Criminisi, “Signle view metrology: Algorithms and Applications”, 1999.

Convolutional Neural Networks for Feature Extraction



- This is the vgg-f CNN architecture
- Trained to classify 1000 object classes

Pedestrian Detection Project

- We have two classes i.e. pedestrian and non-pedestrian
- The CNN is trained to classify 1000 object classes and pedestrian may not be one of them
- Training a CNN from scratch for pedestrian detection requires a very large training dataset.
- How can we use a pre-trained CNN for our task?
- We can use a pre-trained CNN to extract image features i.e. output of any feature map (generally the last layer)
- And train another classifier to differentiate pedestrians from non-pedestrians

Demonstration of using CNN

- Matconvnet installation
- Download a pre-trained CNN model
- There are many options to chose from
- Let us try it out to classify some images from the Internet
- Now let us look at the feature maps

Summary

- Vanishing points and lines
- Projective invariants
- Cross-ratio
- Measuring the height of objects from a single image
- Homography
- Image rectification

Acknowledgements: Material for this lecture was taken from Criminisi’s papers, Wikipedia and previous lectures delivered by Du Huynh and Peter Kovesi.