universität innsbruck

Lecture 8. Stereopsis II

703142. Computer Vision

Assoz.Prof. Antonio Rodríguez-Sánchez, PhD.

Source: Hartley and Zisserman, 2004

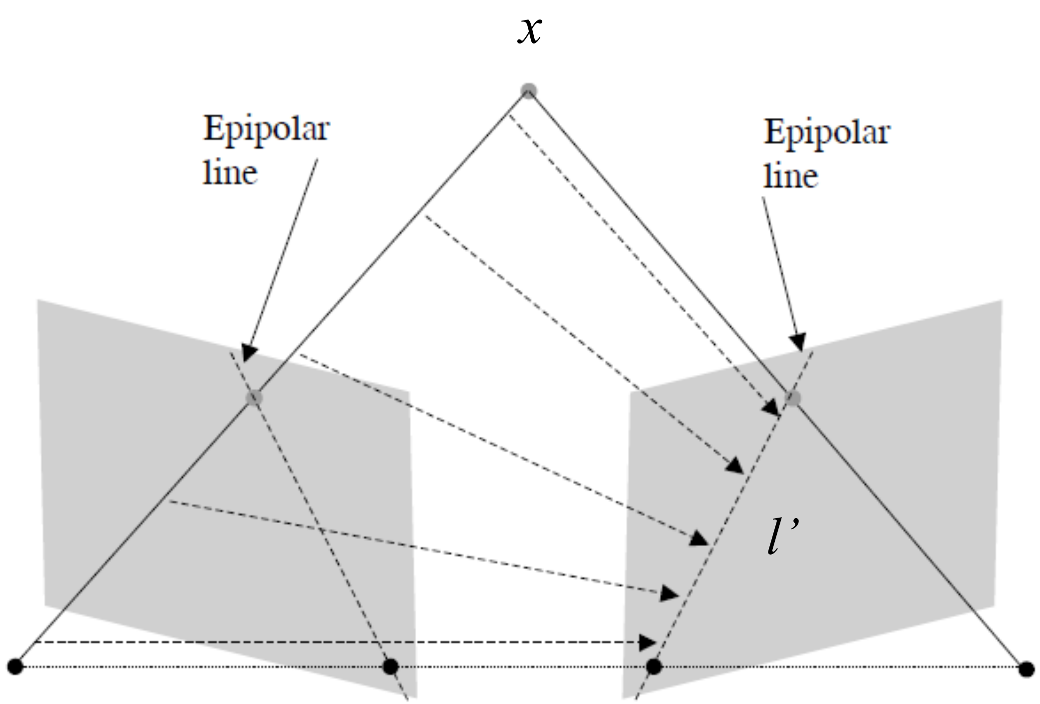
Outline

- Introduction
- 3D shapes from 2D images
- Stereo vision
- Correspondence
- Epipolar geometry
- The fundamental matrix
- The essential matrix
- RANSAC
- 3D reconstruction

The fundamental matrix

- Algebraic representation of epipolar geometry
- Extracted from the mapping between a point and its epipolar line
- There is a map

$$x \mapsto l'$$

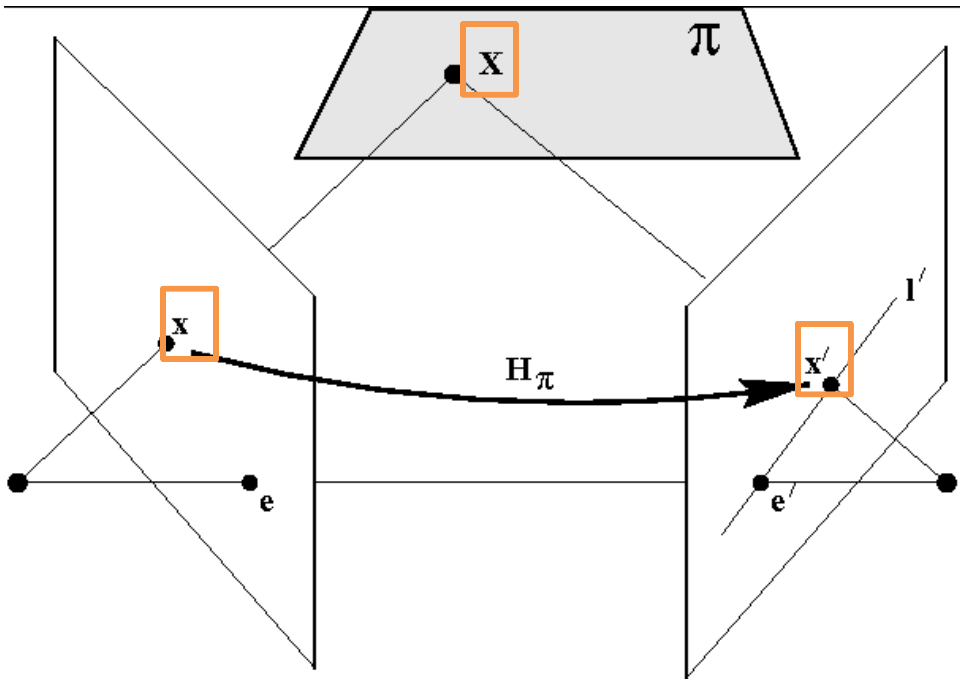


The fundamental matrix

- There is a map

$$x \mapsto l'$$

- Consider point $\mathbf{X}$
  - Projection is $\mathbf{x}$ on the left image plane
  - Projection is $\mathbf{x}_i'$ on the right image plane

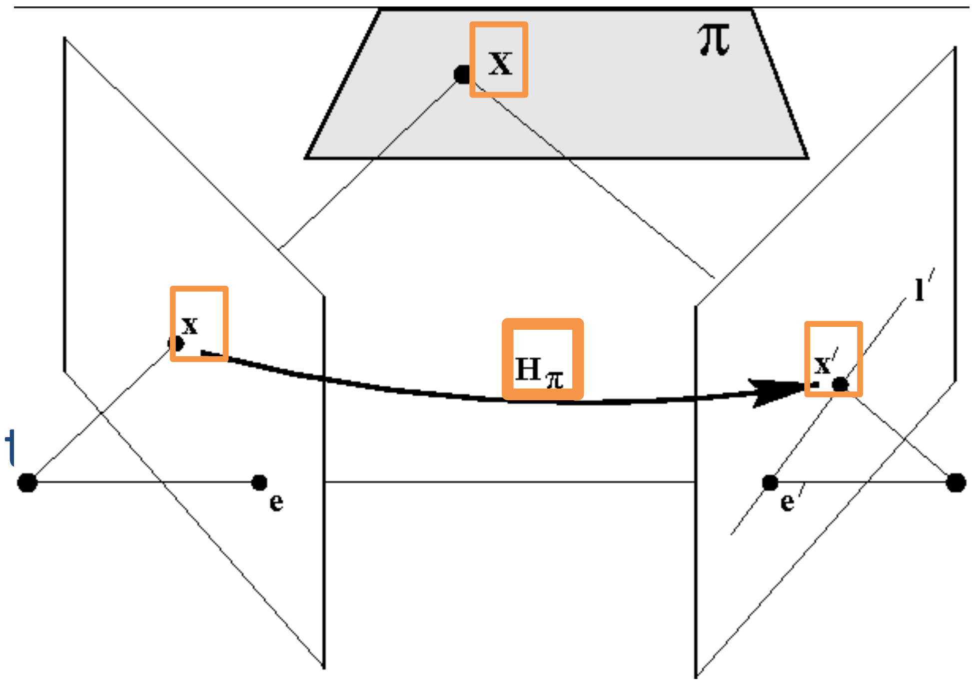


The fundamental matrix

- There is a map

$$x \mapsto l'$$

- Consider point $\mathbf{X}$
  - Projection is $\mathbf{x}$ on the left image plane
  - Projection is $\mathbf{x}_i'$ on the right image plane
- Then there is a 2D homography $\mathbf{H}_\pi$ mapping each $\mathbf{x}$ into $\mathbf{x}_i'$

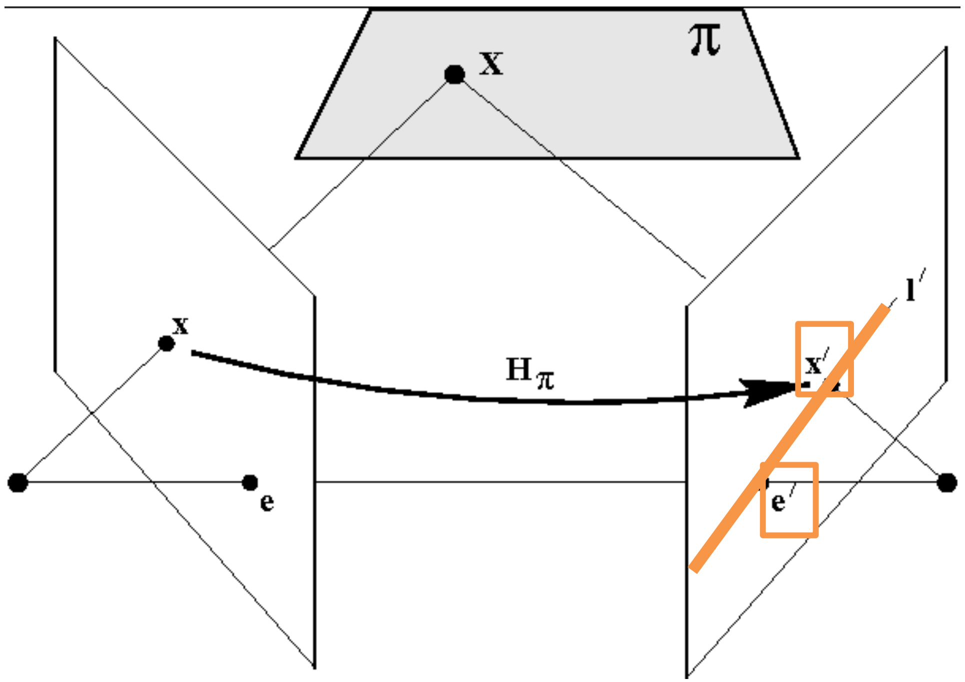


The fundamental matrix

- There is a 2D homography $\mathbf{H}_\pi$ mapping each $\mathbf{x}$ into $\mathbf{x}_i'$
- The epipolar line is the line that goes through $\mathbf{x}'$ and $\mathbf{e}'$

$$l' = e' \times x'$$

crossp



Note: Skew symmetric matrices and cross products

- A vector $\mathbf{a}=(a_1,a_2,a_3)^T$ can be written as a skew matrix as

$$[\mathrm{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

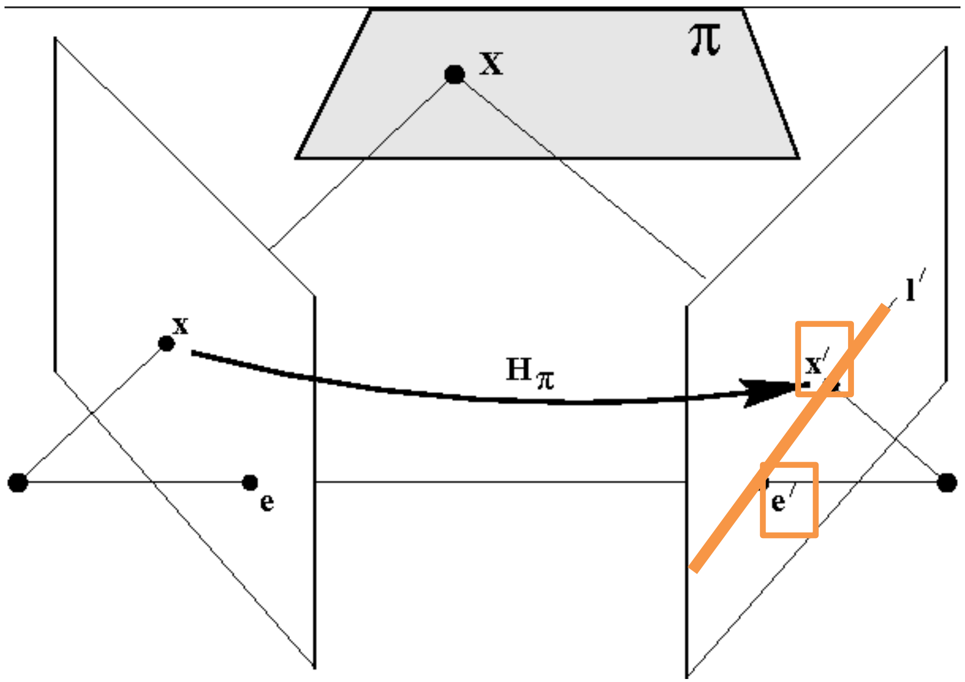
- The cross product is related to skew matrices according to

$$\mathrm{a} \times \mathrm{b} = [\mathrm{a}]_\times \mathrm{b} = (\mathrm{a}^T [\mathrm{b}])^T$$

The fundamental matrix

- There is a 2D homography $\mathbf{H}_\pi$ mapping each $\mathbf{x}$ into $\mathbf{x}_i'$
- The epipolar line is the line that goes through $\mathbf{x}'$ and $\mathbf{e}'$

$$l' = e' \times x'$$



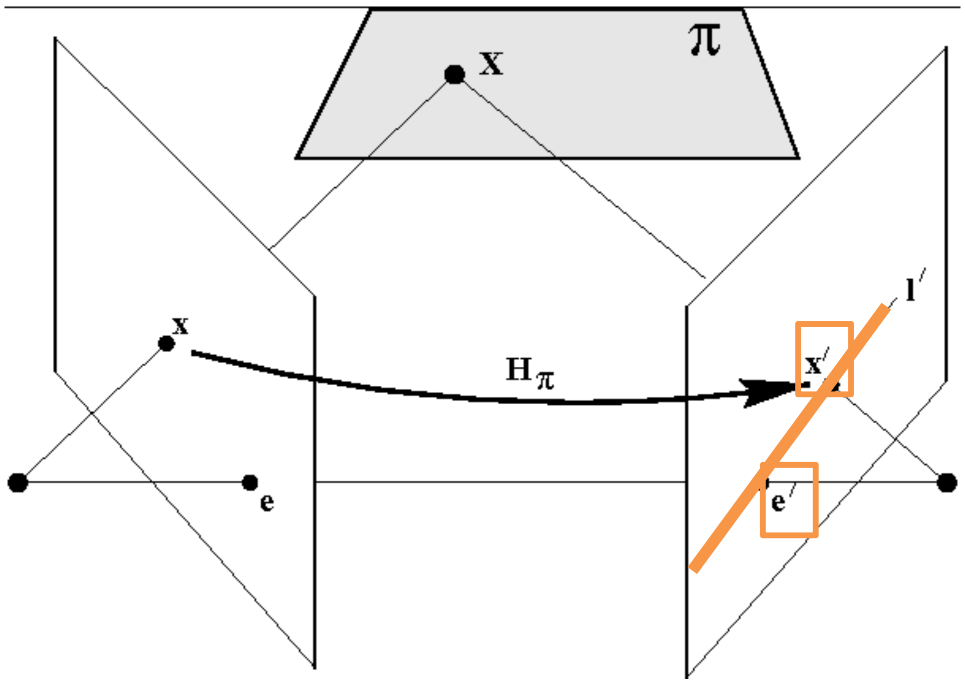
The fundamental matrix

- There is a 2D homography $\mathbf{H}_\pi$ mapping each $\mathbf{x}$ into $\mathbf{x}_i'$
- The epipolar line is the line that goes through $\mathbf{x}'$ and $\mathbf{e}'$

$$l' = e' \times x'$$

- But $\mathbf{x}' = \mathbf{H}_\pi\ \mathbf{x}$, so we have

$$l' = e' \times x' = \boxed{[e']_\times H_\pi} x = Fx$$



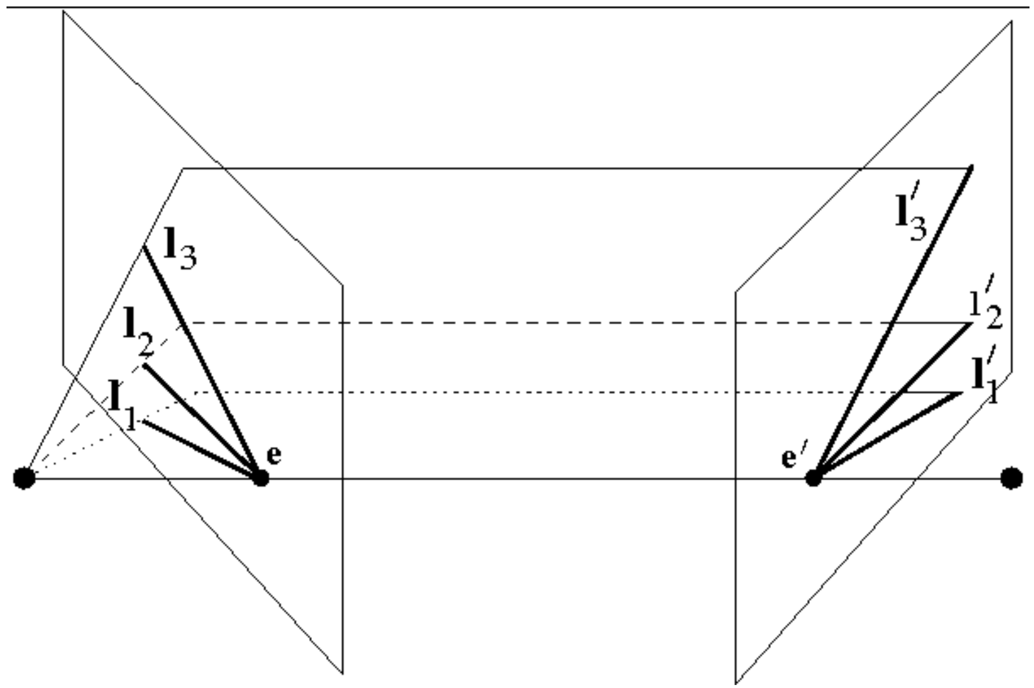
F .. Fundamental Mat

The fundamental matrix

- The fundamental matrix is

$$F = [e']_{\times} H_{\pi}$$

- ***F*** represents a mapping from the 2-dimensional projective plane $\mathbf{P}^2$ of the first image to the pencil of epipolar lines through $\mathbf{e}'$
  - That is, it is a mapping from 2D into 1D



The fundamental matrix

- The fundamental matrix is $F = [e']_\times H_\pi$
- For all corresponding points $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$, The fundamental matrix satisfies

$$x'^T F x = 0$$

The fundamental matrix

- The fundamental matrix is $F = [e']_\times H_\pi$
- For all corresponding points $x \leftrightarrow x'$, The fundamental matrix satisfies

$$x'^T F x = 0$$

- The essential matrix is the specific case where it is assumed that cameras are calibrated
  - Thus, the fundamental matrix is the generalized essential matrix

Outline

- Introduction
- 3D shapes from 2D images
- Stereo vision
- Correspondence
- Epipolar geometry
- The fundamental matrix
- The essential matrix
- RANSAC
- 3D reconstruction

Camera models

- Intrinsic parameters
  - Relate camera's coordinate system to idealized coordinate system
  - That is, link the pixel coordinates of an image point with corresponding coordinates in the camera reference frame.
- Extrinsic parameters
  - Relate the camera's coordinate system to a fixed world coordinate system and specify its position and orientation in space
- The problem of estimating the intrinsic and extrinsic parameters of a camera is known as geometric camera calibration

Geometry of image formation

- Intrinsic parameters
  - Camera to pixel transformation
  - Typically given by 2 sets of parameters
    1. The focal length, f, serving to capture the (perspective) projection.
    2. The pixel coordinates of the principle point (image center), (o_x, o_y), and the effective pixel horizontal and vertical dimensions, (s_x, s_y), serving to capture the transformation between camera frame coordinates and pixel coordinates.
  - We already have considered how to incorporate the focal length.
  - Letting (x_i, y_i) be the pixel coordinates, we incorporate the second set of parameters via

$$x=-(x_i-o_x)s_x$$

$$y=-(y_i-o_y)s_y$$

  - There are 5 total intrinsic parameters: (f, o_x, o_y, s_x, s_y)

Note

- Extrinsic parameters
  - Camera to world transformation
  - Typically given via two sets of parameters
    1. A 3D translation vector, ***T***, describing the relative positions of the two reference frames.
    2. A 3x3 rotation matrix, ***R***, that brings the corresponding axes of the two frames into alignment.
  - Letting P_c and P_w be the camera and world coordinates of the same point, we write:

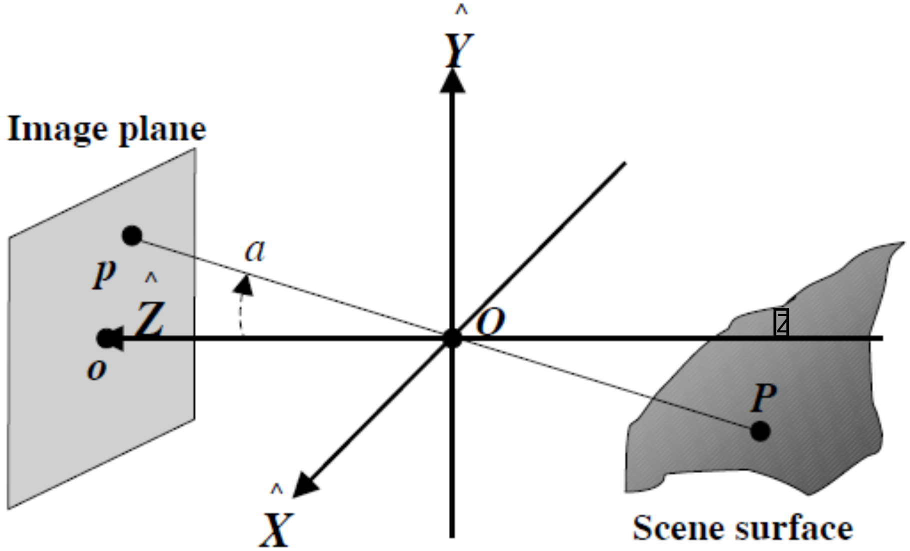
$$P_c = R(P_w - T)$$

  - By definition, ***R***, is completely specified be 3 parameters (e.g., rotation about each of the coordinate axes); so, there are 6 extrinsic parameters in total (3 for ***T***; 3 for ***R***).

Source: Richard Wildes slides

Geometry of image formation

- Fundamental projections: **Perspective**
  - The length of ***P*** is $\|P\| = -Z\sec a = -(P\cdot\hat{Z})\sec a$
  - The length of ***p*** is $\|p\| = f\sec a$
  - So $$\frac{1}{f}p = \frac{1}{P\cdot\hat{Z}}P$$
  - In component form $$\frac{x}{f} = \frac{X}{Z}, \frac{y}{f} = \frac{Y}{Z}$$
  - Or $$x = f\frac{X}{Z}, y = f\frac{Y}{Z}$$



Source: Richard Wildes slides

Geometry of image formation

- World-image transformation
  - Recall the camera frame expression of perspective projection

$$x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$$

- We substitute the intrinsic parameterization on the lhs and the extrinsic parameterization on the rhs to find that (with $\boldsymbol{R_i}$ the i-th row of **R**)

$$-(x_i - o_x)s_x = f\frac{R_1^T(P_w - T)}{R_3^T(P_w - T)}$$

$$-(y_i - o_y)s_y = f\frac{R_2^T(P_w - T)}{R_3^T(P_w - T)}$$

P_c = R * (P_w - T)

- **Which relates the 3D coordinates of a point in the world frame to its corresponding image coordinates.**

Geometry of image formation

- World-image transformation
  - Rewrite in matrix formulation

$$-(x_i - o_x)s_x = f\frac{R_1^T(P_w - T)}{R_3^T(P_w - T)}$$

$$-(y_i - o_y)s_y = f\frac{R_2^T(P_w - T)}{R_3^T(P_w - T)}$$

  - Define matrices that encapsulate intrinsic, $\mathbf{M}_{\text{int}}$, and extrinsic, $\mathbf{M}_{ext}$ parameters

$$\mathbf{M}_{\text{int}} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{M}_{ext} = \begin{pmatrix} \bullet & \bullet & \bullet & -R_1^T T \\ \bullet & \mathbf{R} & \bullet & -R_2^T T \\ \bullet & \bullet & \bullet & -R_3^T T \end{pmatrix}$$

Geometry of image formation

- World-image transformation

$$\mathbf{M}_{\text{int}} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{M}_{ext} = \begin{pmatrix} \bullet & \bullet & \bullet & -R_1^T T \\ \bullet & \mathbf{R} & \bullet & -R_2^T T \\ \bullet & \bullet & \bullet & -R_3^T T \end{pmatrix}$$

  – And concatenate them to write

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{M}_{\text{int}} \mathbf{M}_{ext} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Remark: $(x_i, y_i) = (x_1/x_3, x_2/x_3)$

The essential matrix

- The fundamental matrix is $F = [e']_\times H_\pi$
- For all corresponding points $x \leftrightarrow x'$, The fundamental matrix satisfies

$$x'^T F x = 0$$

- The essential matrix is the specific case where it is assumed that the cameras are normalized
  - Thus, the fundamental matrix is the generalized essential matrix

The essential matrix

- The fundamental matrix is $F = [e']_{\times} H_{\pi}$
- For all corresponding points $x \leftrightarrow x'$, The fundamental matrix satisfies

$$x'^{T} F x = 0$$

- The essential matrix is the specific case where it is assumed that the cameras are normalized
  - Thus, the fundamental matrix is the generalized essential matrix
  - We assume that the world coordinate system is centered on the first camera, extrinsic parameters are $\{\mathbf{I},\mathbf{0}\}$. $\mathbf{M}_{int1} = \mathbf{I}$
  - The second camera may be in any general position $\{\Omega, \tau\}$. $\mathbf{M}_{int2} = \mathbf{I}$

The essential matrix

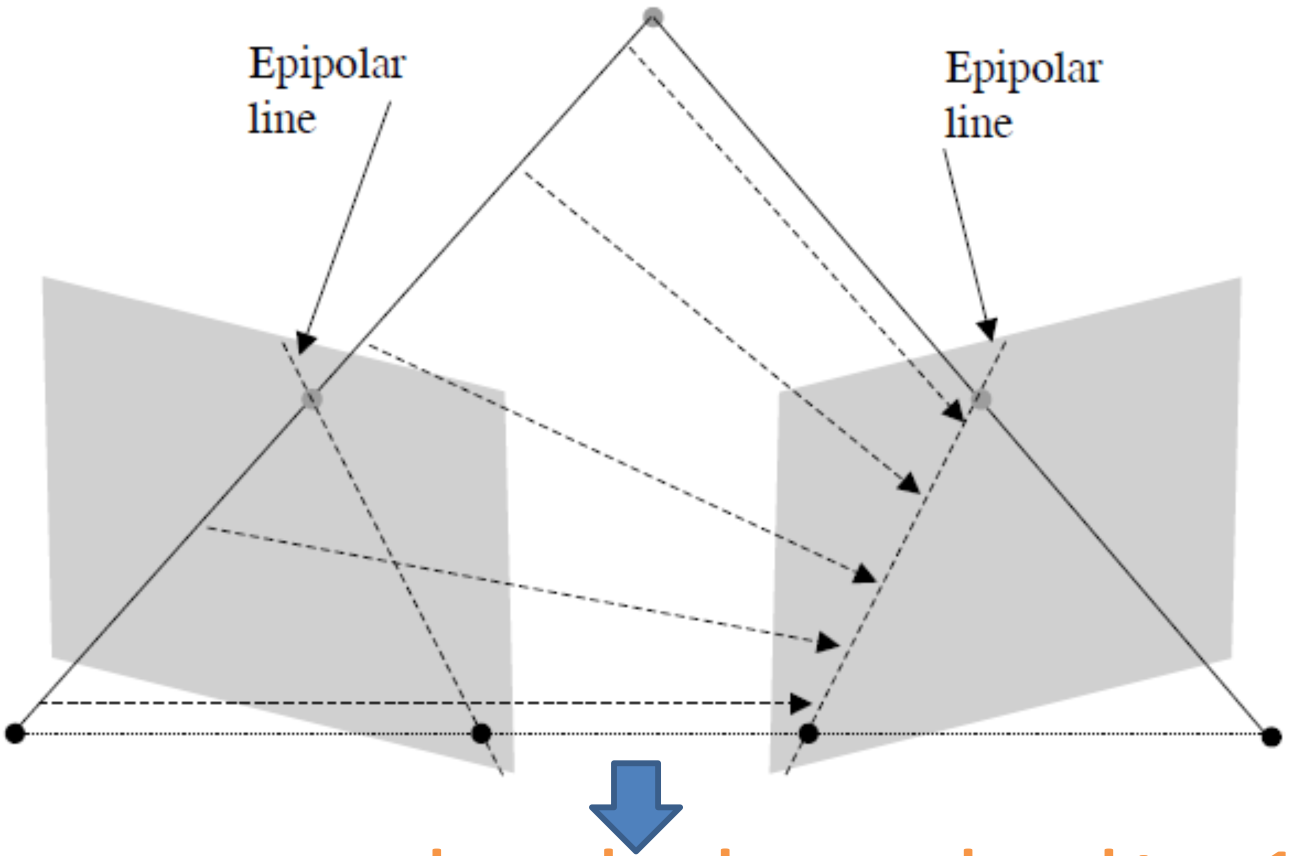
- For all corresponding points in homogeneous coordinates. The essential matrix satisfies

$$\tilde{\mathbf{x}}_2^{\mathbf{T}}\mathbf{E}\tilde{\mathbf{x}}_1 = \mathbf{0}$$

- The essential matrix is the specific case where it is assumed that the cameras are normalized
  - Thus, the fundamental matrix is the generalized essential matrix
  - We assume that the world coordinate system is centered on the first camera, extrinsic parameters are $\{\mathbf{I},\mathbf{0}\}$. $\mathbf{M}_{int1} = \mathbf{I}$
  - The second camera may be in any general position $\{\Omega, \tau\}$. $\mathbf{M}_{int2} = \mathbf{I}$

Epipolar geometry

- Given a stereo pair of cameras and a point in 3D space
  - There is a plane that goes through the point and the centres of projection of the cameras
    - We call this plane the epipolar plane
  - The lines where the plane intersects the images are called conjugate epipolar lines



Stereo correspondence has been reduced to a 1D search !

The essential matrix

- For all corresponding points in homogeneous coordinates. The essential matrix satisfies

$$\tilde{\mathbf{x}}_{\mathbf{2}}^{\mathbf{T}}\mathbf{E}\tilde{\mathbf{x}}_{\mathbf{1}} = \mathbf{0}$$

- The epipolar lines are easily retrieved from the essential matrix. The condition for a point being on a line

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0.$$

The essential matrix

- For all corresponding points in homogeneous coordinates. The essential matrix satisfies

$$\tilde{\mathbf{x}}_2^{\mathbf{T}} \mathbf{E} \tilde{\mathbf{x}}_1 = \mathbf{0}$$

- The epipolar lines are easily retrieved from the essential matrix. The condition for a point being on a line

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0.$$

- The line $\mathbf{l_1} = \tilde{\mathbf{x}}_2 \mathbf{E}$ is the epipolar line in image 1 due to the point $\mathbf{x}_2$ in image 2

The essential matrix

- For all corresponding points in homogeneous coordinates. The essential matrix satisfies

$$\tilde{\mathbf{x}}_2^{\mathbf{T}}\mathbf{E}\tilde{\mathbf{x}}_1 = \mathbf{0}$$

- The epipolar lines are easily retrieved from the essential matrix. The condition for a point being on a line
- The line $\mathbf{l}_1 = \tilde{\mathbf{x}}_2^{\mathbf{T}}\mathbf{E}$ is the epipolar line in image 1 due to the point $\mathbf{x}_2$ in image 2
- Similarly $\mathbf{l}_2 = \tilde{\mathbf{x}}_1^{\mathbf{T}}\mathbf{E}^{\mathbf{T}}$

The essential matrix

- For all corresponding points in homogeneous coordinates. The essential matrix satisfies

$$\tilde{\mathbf{x}}_2^{\mathbf{T}} \mathbf{E} \tilde{\mathbf{x}}_1 = \mathbf{0}$$

- The epipoles can also be extracted from the essential matrix
  - Every epipolar line in in image 1 passes through the epipole $\tilde{\mathbf{e}}_1$, such that $\tilde{\mathbf{x}}_2^{\mathbf{T}} \mathbf{E} \tilde{\mathbf{e}}_1 = 0$
  - Epipoles can be retrieved by computing the singular value decomposition $\mathbf{E} = \mathbf{U} \mathbf{L} \mathbf{V}^T$ of the essential matrix
  - Setting $\tilde{\mathbf{e}}_1$ to the last column of $\mathbf{V}$ and
  - $\tilde{\mathbf{e}}_2$ to the last row of $\mathbf{U}$

The essential matrix

- The essential matrix is defined

$$\mathbf{E} = \tau_{\times}\mathbf{\Omega}$$

  - Applying $\mathbf{E} = \mathbf{ULV}^T$
  - $\tau_{\times} = \mathbf{ULWU}^T$
  - $\mathbf{\Omega} = \mathbf{UW}^{-1}\mathbf{V}^T$

$$\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3D reconstruction

- Goal
  - Given binocular stereo images and their correspondences.
  - Recover the 3D geometry of the imaged scene.
- What can be achieved
  - Given intrinsic and extrinsic geometry of the cameras: Absolute Euclidean reconstruction.
  - Given only intrinsic camera geometries: Reconstruction up to a scale factor.
  - Given no information on camera geometries: Reconstruction up to a projective transformation.

Outline

- Introduction
- 3D shapes from 2D images
- Stereo vision
- Correspondence
- Epipolar geometry
- The fundamental matrix
- The essential matrix
- RANSAC
- 3D reconstruction

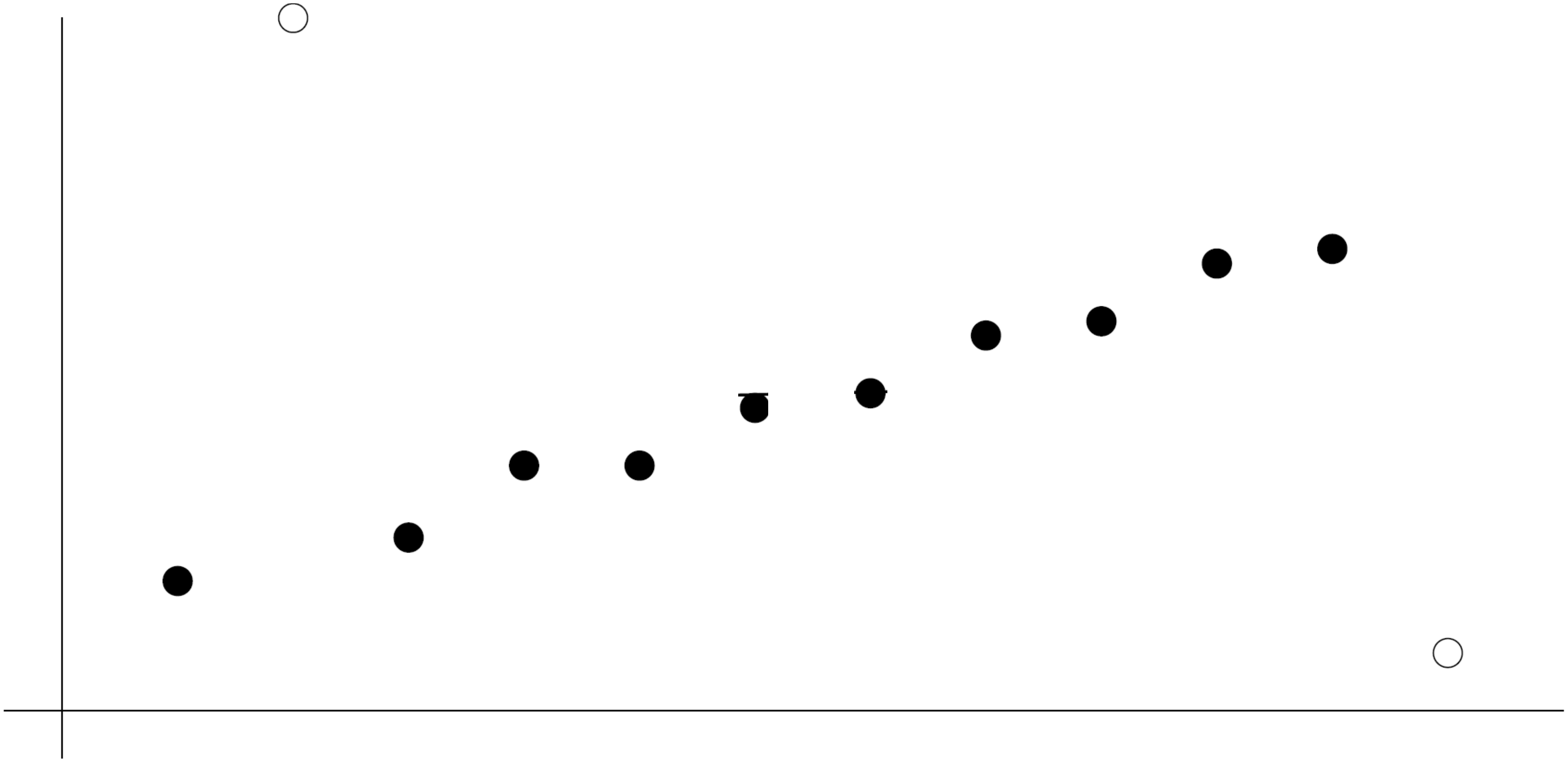
RANSAC

- RAndom SAmple Consensus
- Fischler and Bolles, 1981
  - Estimating a straight line
  - Fit of a set of 2-dimensional points
  - Equivalent to estimating a 1-dimensional affine transformation

$$x' = ax + b$$

RANSAC

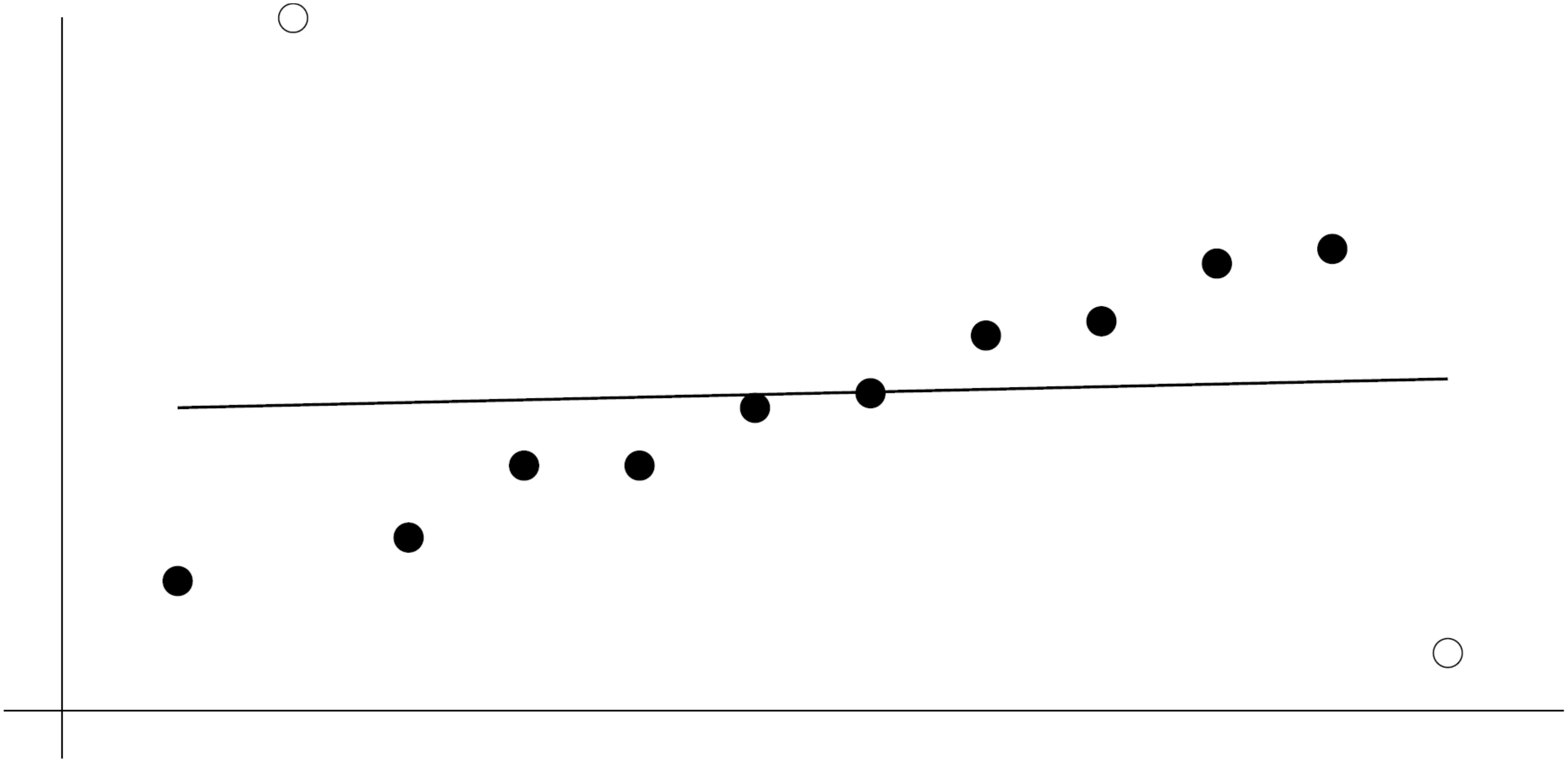
- RAndom SAmple Consensus



Source: Hartley and Zisserman, 2004

RANSAC

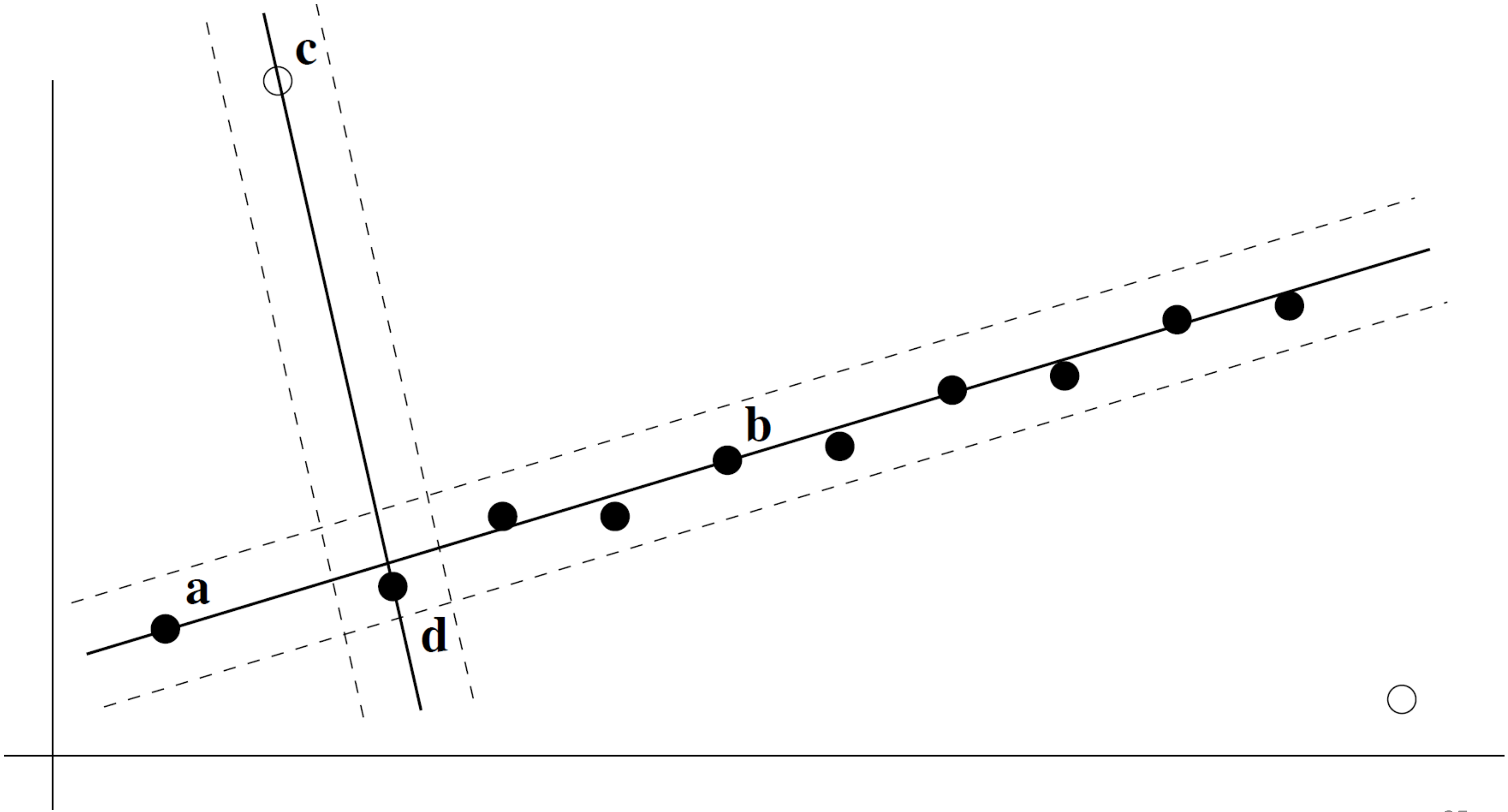
- RAndom SAmple Consensus



Source: Hartley and Zisserman, 2004

RANSAC

- RAndom SAmple Consensus



Source: Hartley and Zisserman, 2004

RANSAC

- RAndom SAmple Consensus

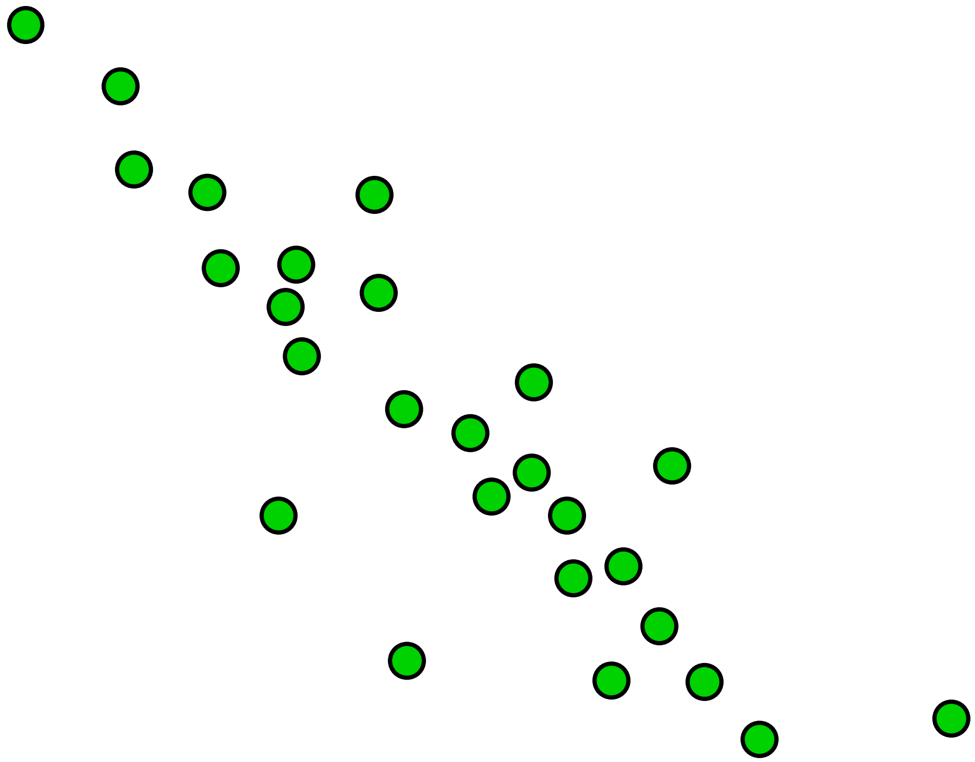
1. Repeat k times:
    1. Draw a random sample of n data points.
    2. Determine the model parameters using these points.
    3. For each data point outside the sample:
        1. If its distance to the model is less than d, then label it as a “good” point.
    4. If there are at least t “good” data points, then declare this model “good” and quit this loop.
2. Refit the model with the most “good” points using all of them, including the n sample points.

RANSAC

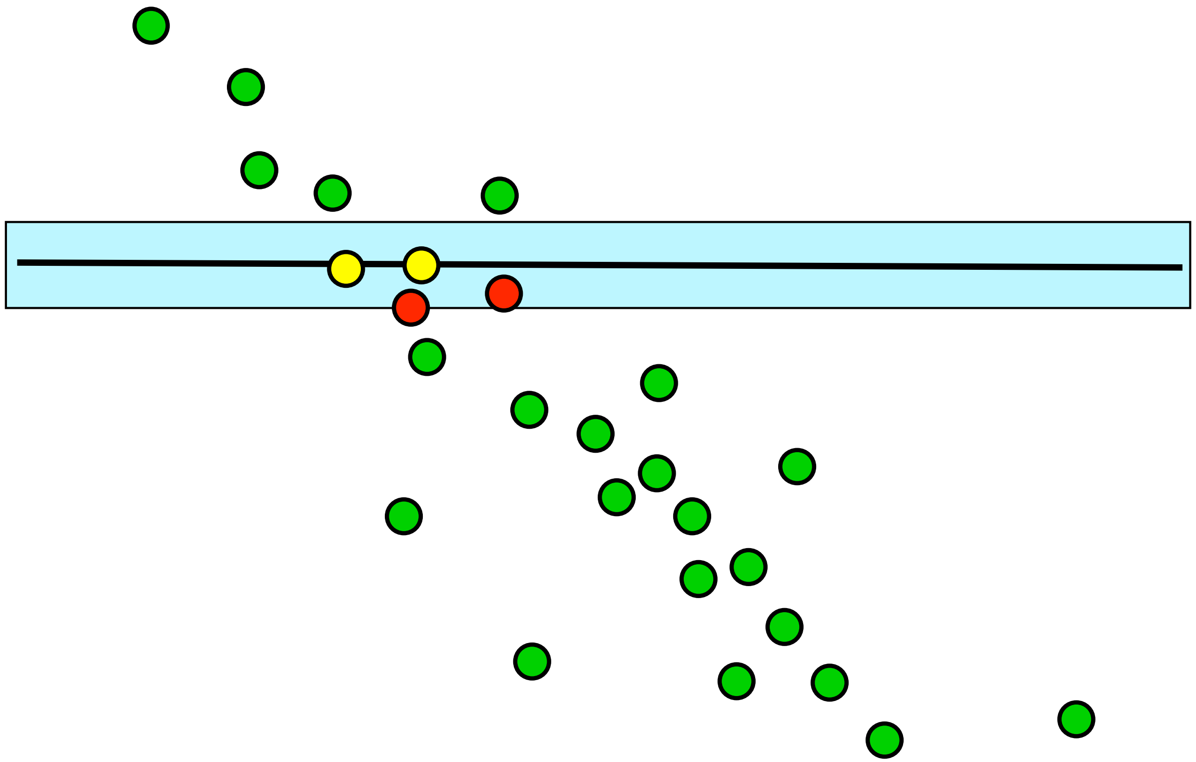
- RAndom SAmple Consensus

1. Repeat $\boxed{k}$ times:
    1. Draw a random sample of $\boxed{n}$ data points.
    2. Determine the model parameters using these points.
    3. For each data point outside the sample:
        1. If its distance to the model is less than $\boxed{d}$, then label it as a “good” point.
    4. If there are at least $\boxed{t}$ “good” data points, then declare this model “good” and quit this loop.
2. Refit the model with the most “good” points using all of them, including the n sample points.

RANSAC

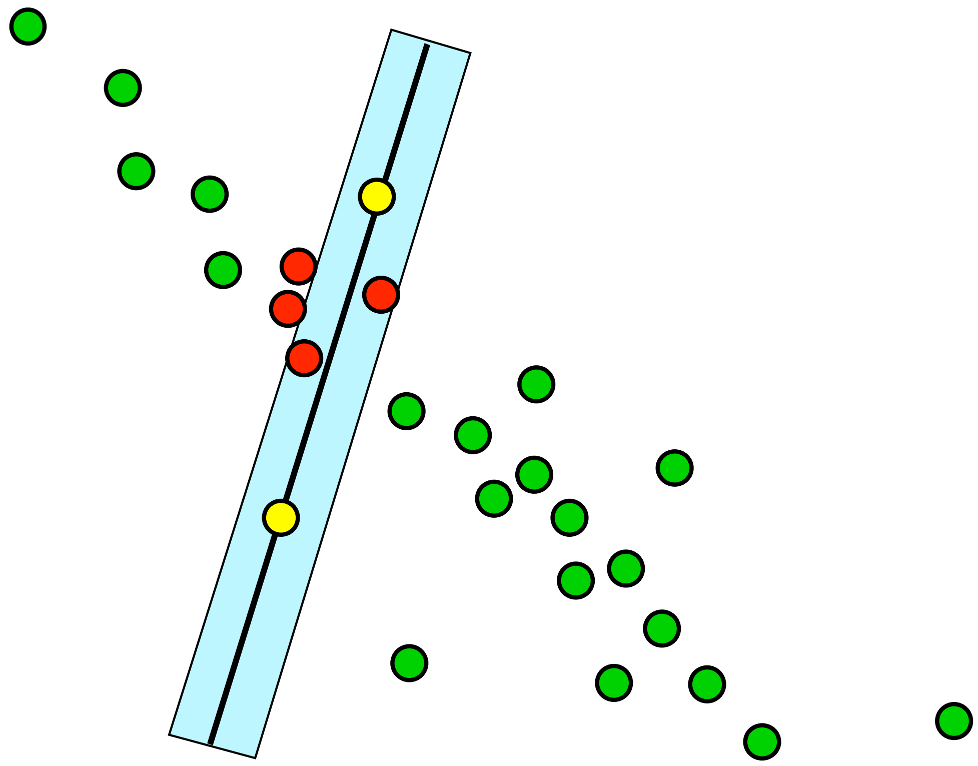


RANSAC



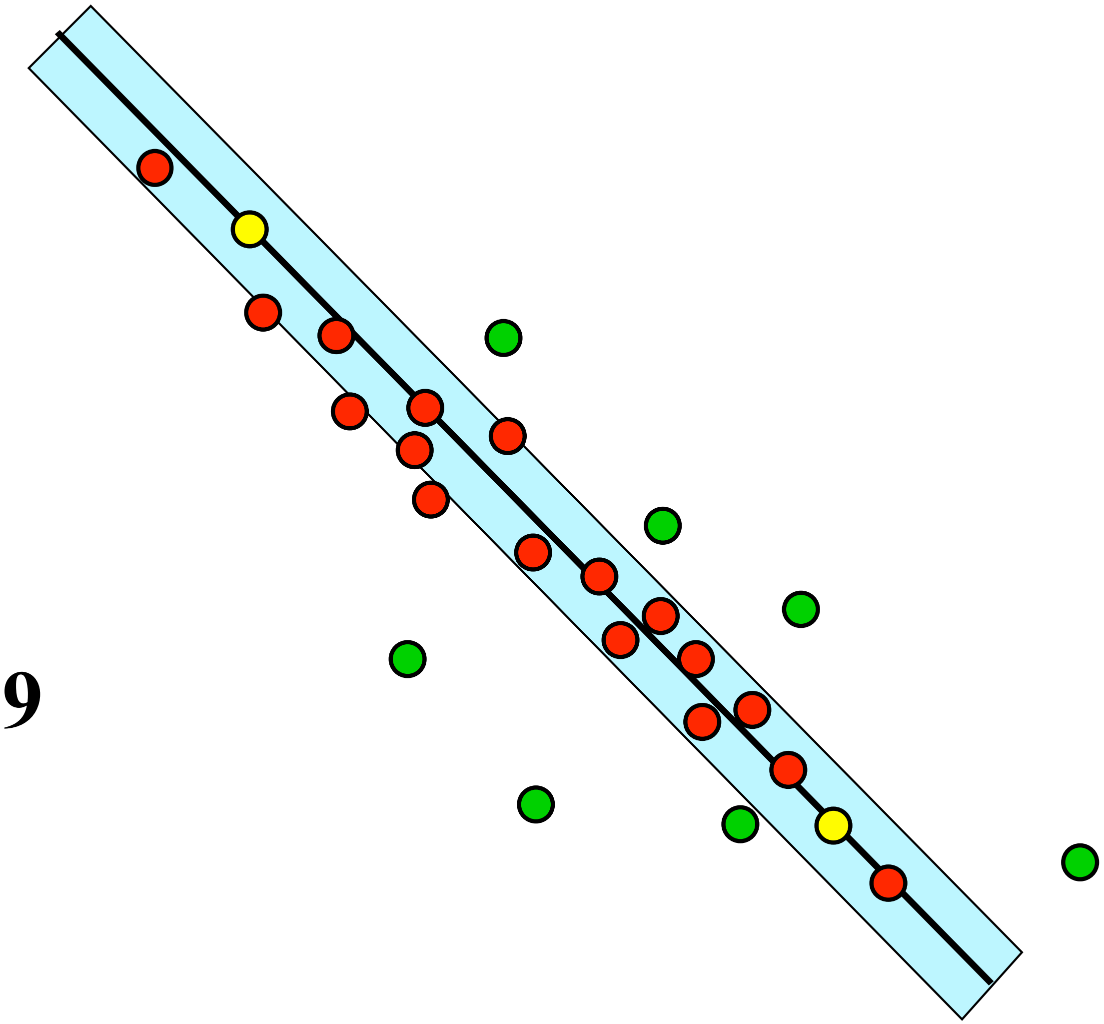
Count = 4

RANSAC



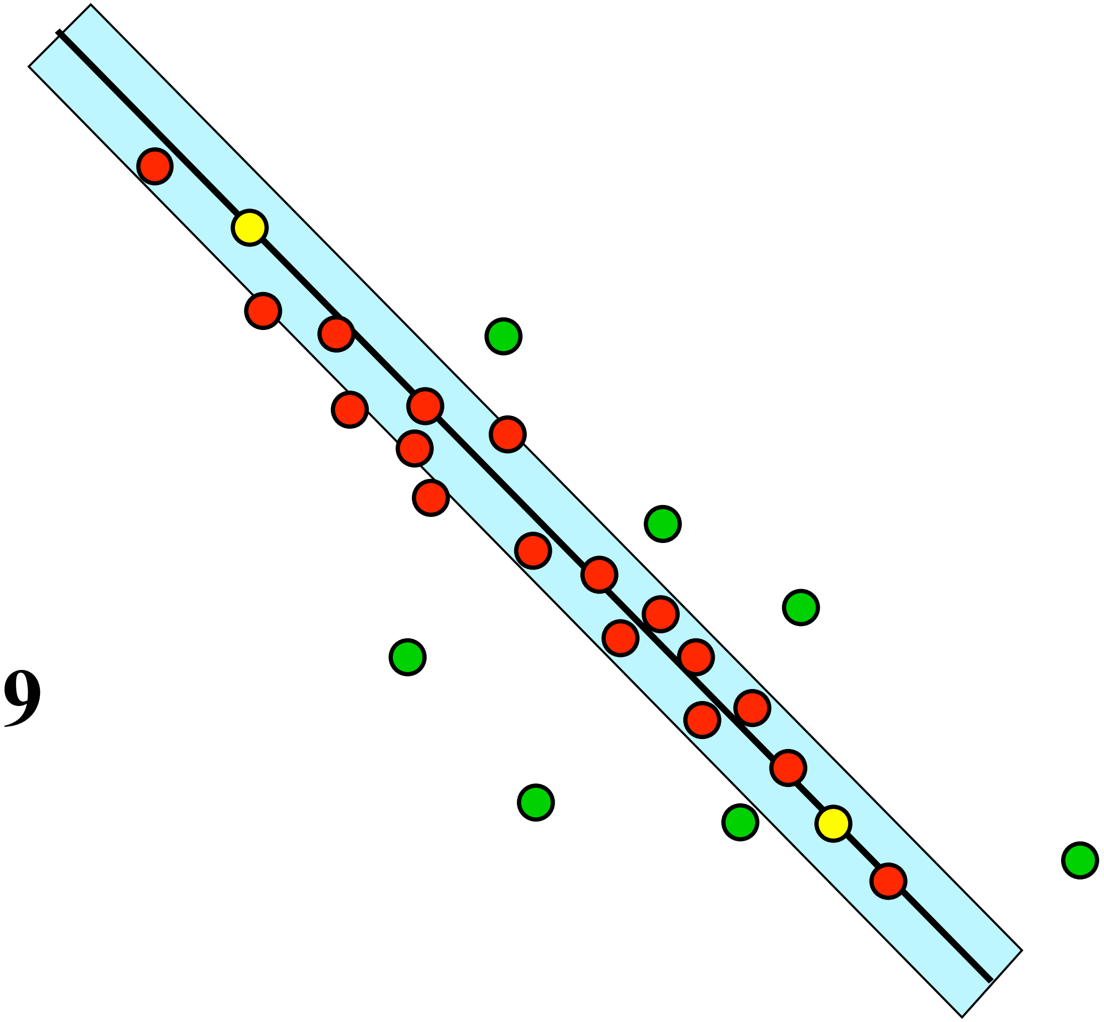
Count = 6

RANSAC



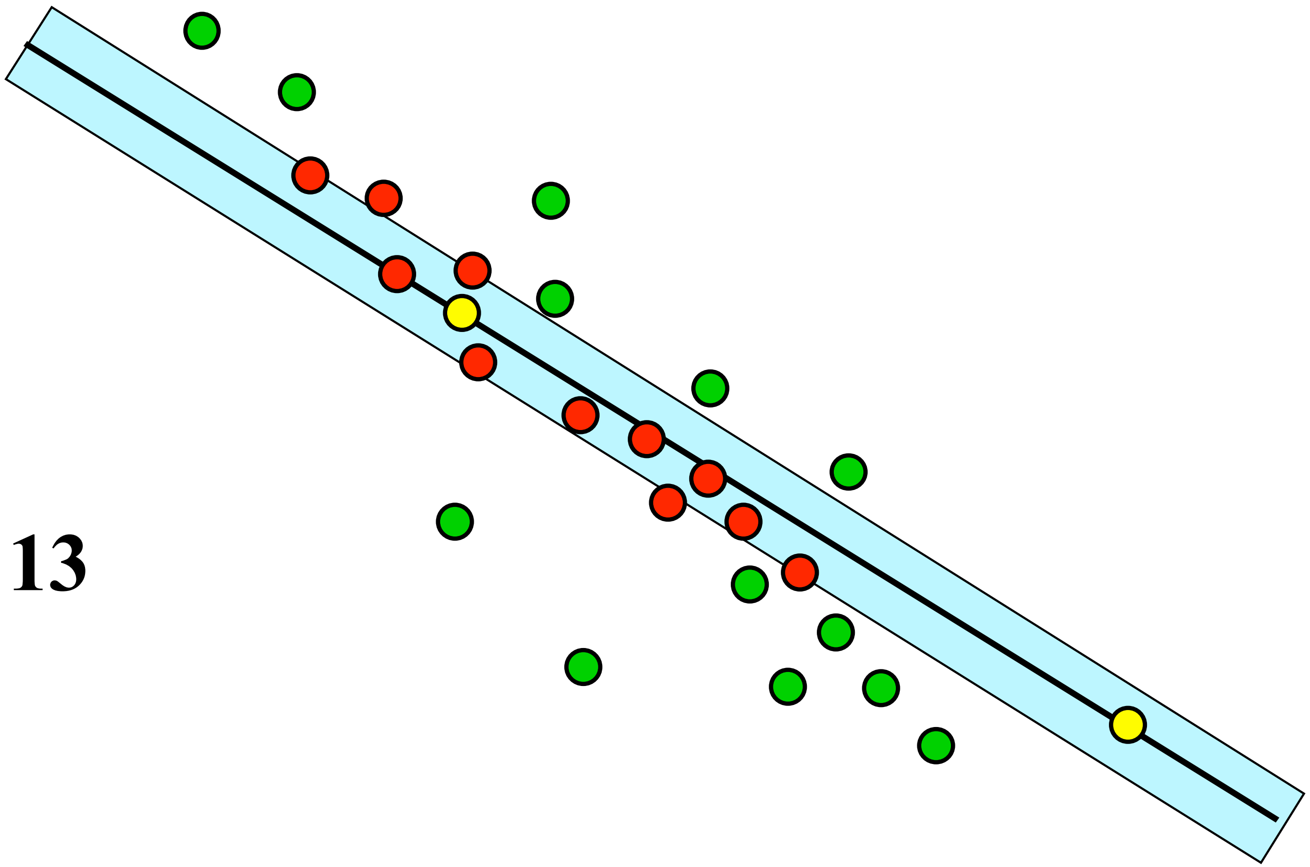
Count = 19

RANSAC



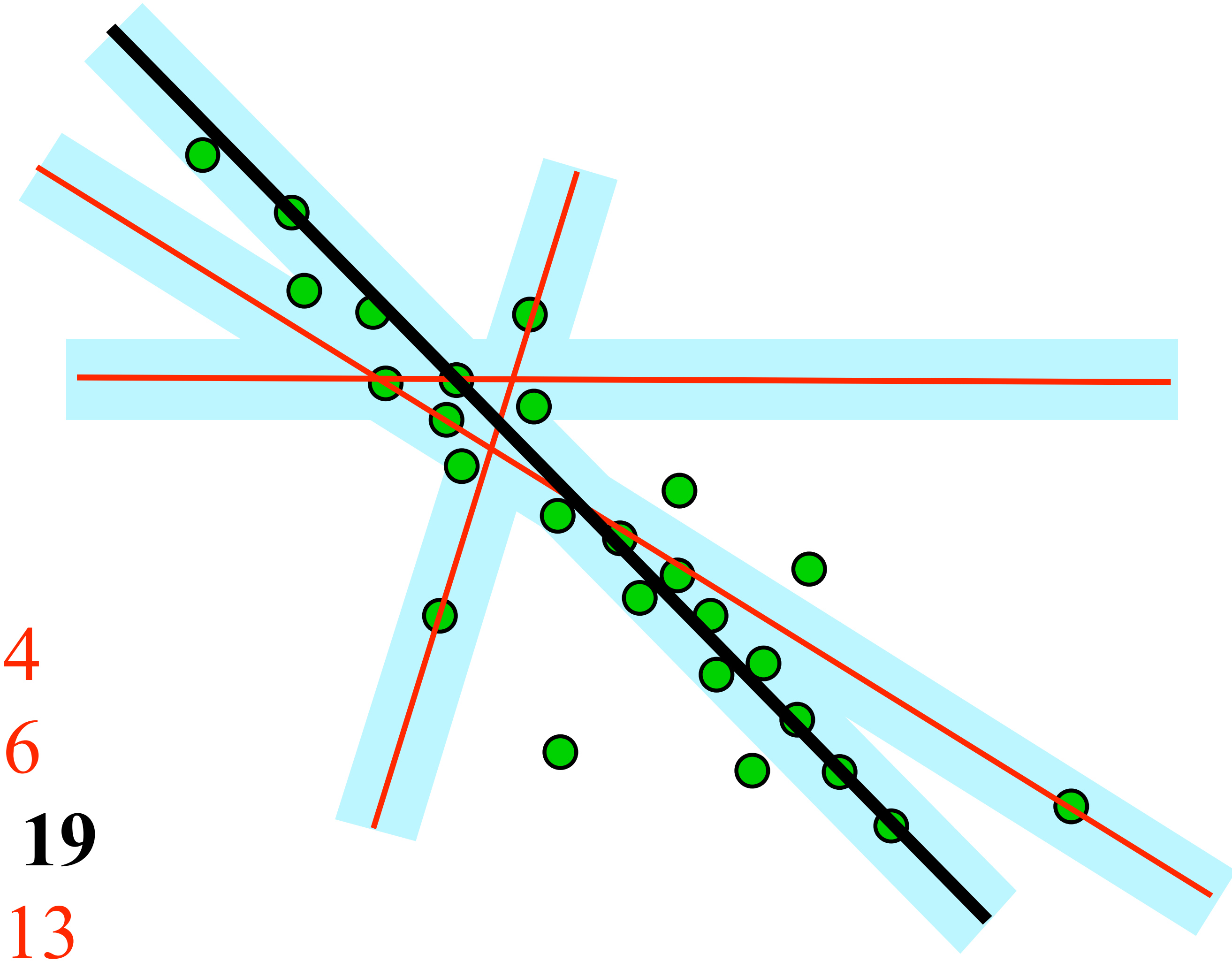
Count = 19

RANSAC



Count = 13

RANSAC



Count = 4

Count = 6

Count = 19

Count = 13

RANSAC

- RAndom SAmple Consensus
- Issues
  - Each sample comprises exactly n data points for an n-parameter model.
  - Measure of the distance of a data point to the model.
  - Choice of the three parameters k, d, and t.

RANSAC

- k: How many trials?
- It is unfeasible and unnecessary to try every possible trial
- The number of trials k must be sufficiently high to ensure with probability p that at least one of the random samples of n is free from outliers

RANSAC

- k: How many trials?
- The number of trials k must be sufficiently high to ensure with probability p that at least one of the random samples of n is free from outliers

$$E(k) = P(one\_good\_sample\_in\_one\_draw)$$
$$+ P(one\_good\_sample\_in\_two\_draws) + \ldots$$

RANSAC

- k: How many trials?
- The number of trials k must be sufficiently high to ensure with probability p that at least one of the random samples of n is free from outliers

$$E(k) = P(one\_good\_sample\_in\_one\_draw)$$
$$+ P(one\_good\_sample\_in\_two\_draws) + \ldots$$

- Let's call w the probability of inliers, that is the probability that a selected point is an inlier

$$E(k) = w^n + 2w^n(1-w^n) + 3w^n(1-w^n)^2 + \ldots = w^{-n}$$

RANSAC

- ***k***: How many trials?
- Let's call ***w*** the fraction of inliers, well, the probability that a selected point is an inlier

$$E(k) = w^n + 2w^n(1-w^n) + 3w^n(1-w^n)^2 + \ldots = w^{-n}$$

  - We need to increase confidence, so, select more than ***w***$^{-n}$
  - Add a few standard deviations to ***k***

$$\sigma(k) = \frac{\sqrt{1-w^n}}{w^n}$$

RANSAC

- k: How many trials?
- Let's call w the fraction of inliers, well, the probability that a selected point is an inlier

$$E(k) = w^n + 2w^n(1-w^n) + 3w^n(1-w^n)^2 + \ldots = w^{-n}$$

  - We need to increase confidence, so, select more than w^{-n}
  - Add a few standard deviations to k

$$\sigma(k) = \frac{\sqrt{1-w^n}}{w^n}$$

- The probability z of only outliers $z = (1-w^n)^k$

RANSAC

- k: How many trials?
- Let's call w the fraction of inliers, well, the probability that a selected point is an inlier

$$E(k) = w^n + 2w^n(1-w^n) + 3w^n(1-w^n)^2 + \ldots = w^{-n}$$

  - We need to increase confidence, so, select more than w^{-n}
  - Add a few standard deviations to k

$$\sigma(k) = \frac{\sqrt{1-w^n}}{w^n}$$

- The probability z of only outliers $z = (1-w^n)^k$

$$k = \frac{\log(z)}{\log(1-w^n)}$$

RANSAC

- *k*: How many trials?, $z=0.01$

Number of samples required

Sample size	Proportion of Outliers 1-w						
n	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177

Table 1: *From [Hartley and Zisserman, 2000]*: The number ***k*** of samples required to ensure, with a probability $p = 0.99$, that at least one sample has no outliers for a given sample size s and a proportion of outliers ϵ.

RANSAC

- *d*: How close must a point to be good?
  - It is usually chosen empirically
    - Expected measurement error
    - Eyeballing
    - Trial-and-error

RANSAC

- *d*: How close must a point to be good?
  - It is usually chosen empirically
    - Expected measurement error
    - Eyeballing
    - Trial-and-error
  - But there are fancier methods too

RANSAC

- *d*: How close must a point to be good?
  - It is usually chosen empirically
    - Expected measurement error
    - Eyeballing
    - Trial-and-error
  - But there are fancier methods too
    - Using a cummulative chi-squared distribution
      - Assuming error is Gaussian with zero mean

RANSAC

- *t*: How many good points for a good fit? Number of points that must agree

RANSAC

- t: How many good points for a good fit? Number of points that must agree
  - A good rule of thumb is to require at least as many "good" points as we would expect from the proportion w of inliers present in the data

$$t = wN$$

  - Where N is the total number of data points

RANSAC

- ***t***: How many good points for a good fit? Number of points that must agree $t = wN$
- Estimating the proportion of inliers ***w***
  - Begin with ***w=0***
  - After each RANSAC iteration
    - Choose a sample and count the number of inliers
    - Set ***w***=(number of inliers)/(total number of points)
    - Recompute ***k***

$$k = \frac{\log(z)}{\log(1 - w^n)}$$

RANSAC

- Compute the 2D homography matrix between two images
  - Input: Two images
  - Output: Homography and correspondence of interesting points

RANSAC

- Compute the 2D homography matrix between two images
  - Input: Two images
  - Output: Homography and correspondence of interesting points
- RANSAC for Automatic Computation of a Fundamental Matrix

$$x'^T \boxed{F} x = 0$$

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$
  1. Compute interest points in each image
     - E.g. Use Harris to detect 7 points in each image

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$
  1. Compute interest points in each image
     - E.g. Use Harris to detect 7 points in each image
  2. Putative correspondences
     - Compute a set of interest point matches based on proximity and similarity of their intensity neighborhood

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$
  1. Compute interest points in each image
     - E.g. Use Harris to detect 7 points in each image
  2. Putative correspondences
     - Compute a set of interest point matches based on proximity and similarity of their intensity neighborhood
  3. Putative $\boldsymbol{F}$: Use the minimal algorithm, if there are three solutions, try them all

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$
  1. Compute interest points in each image
     - E.g. Use Harris to detect 7 points in each image
  2. Putative correspondences
     - Compute a set of interest point matches based on proximity and similarity of their intensity neighborhood
  3. Putative $\boldsymbol{F}$: Use the minimal algorithm, if there are three solutions, try them all
  4. Assess the error
     - Use the symmetric transfer error or the Sampson distance

$$\frac{(x'^T Fx)^2}{(Fx)_1^2 + (Fx)_2^2 + (F^T x')_1^2 + (F^T x')_2^2}$$

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix

$$x'^T \boxed{F} x = 0$$

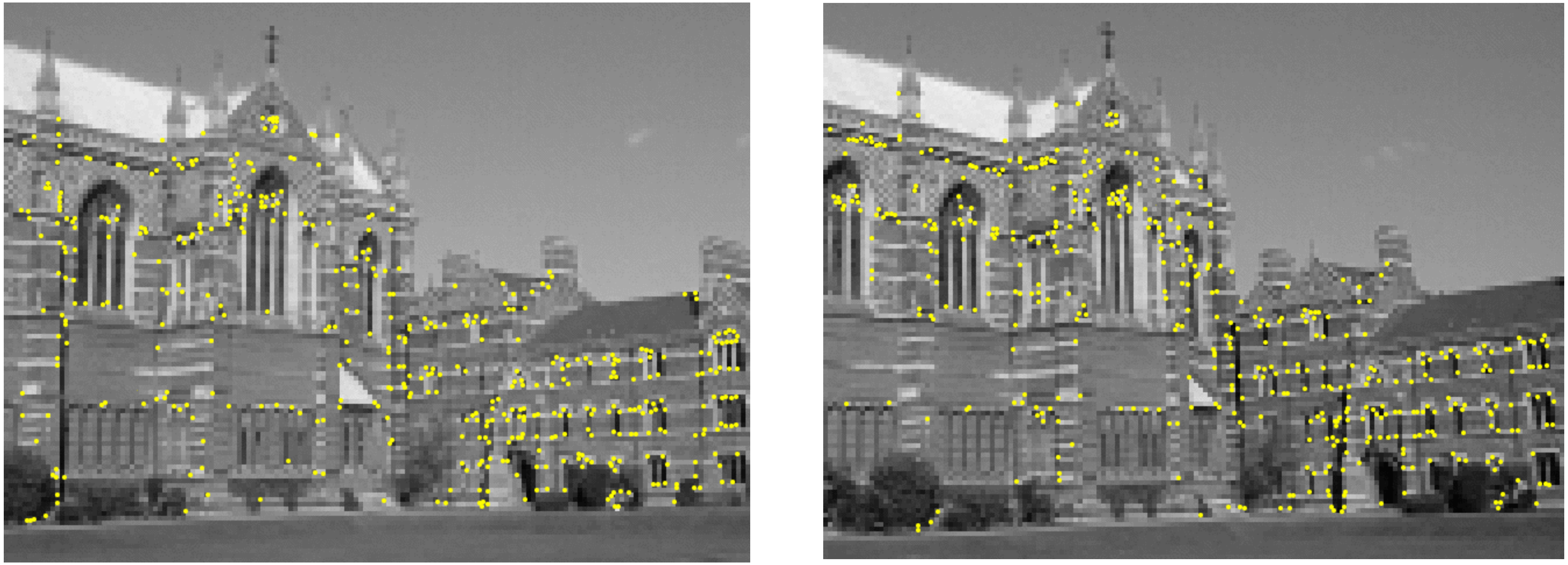
3. Putative $\boldsymbol{F}$: Use the minimal algorithm, if there are three solutions, try them all
4. Assess the error
   - Use the symmetric transfer error or the Sampson distance

$$\frac{(x'^T Fx)^2}{(Fx)_1^2 + (Fx)_2^2 + (F^T x')_1^2 + (F^T x')_2^2}$$

5. Re-estimate $\boldsymbol{F}$ from all inliers of the RANSAC result by optimizing the same error function as used as the distance measure. We may also add new correspondences by searching in a small strip around the epipolar lines.

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^{T}\boxed{F}x = 0$

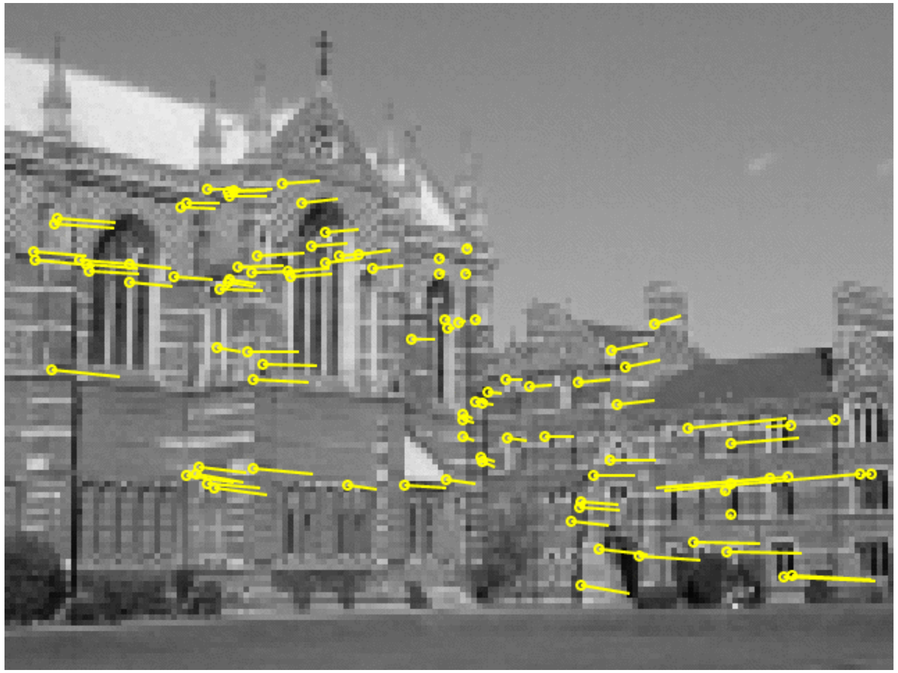


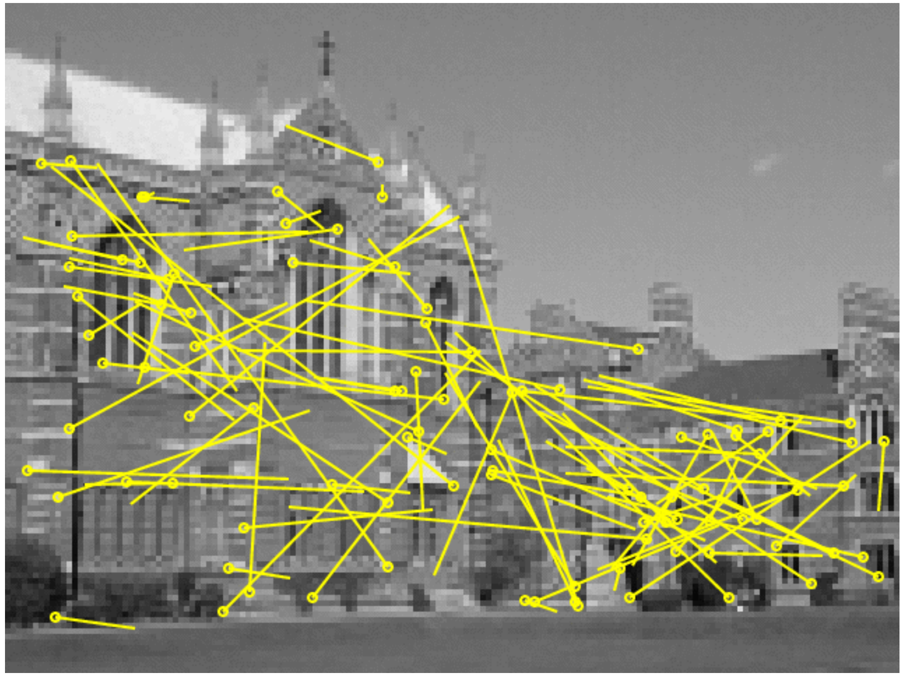
Step 1. ~500 interest points per image

Source: Hartley and Zisserman, 2004

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$



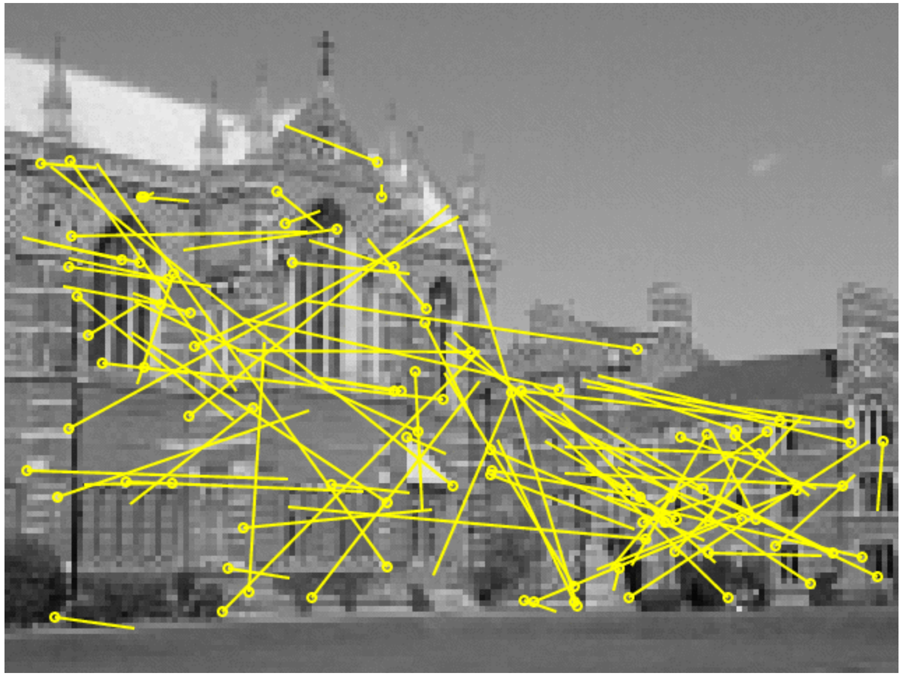


Step 2. Putative matches: 99 inliers, 89 outliers

Source: Hartley and Zisserman, 2004

RANSAC

- RANSAC for Automatic Computation of a Fundamental Matrix $x'^T \boxed{F} x = 0$



Step 5. All 188 putative matches: 157 revised matches

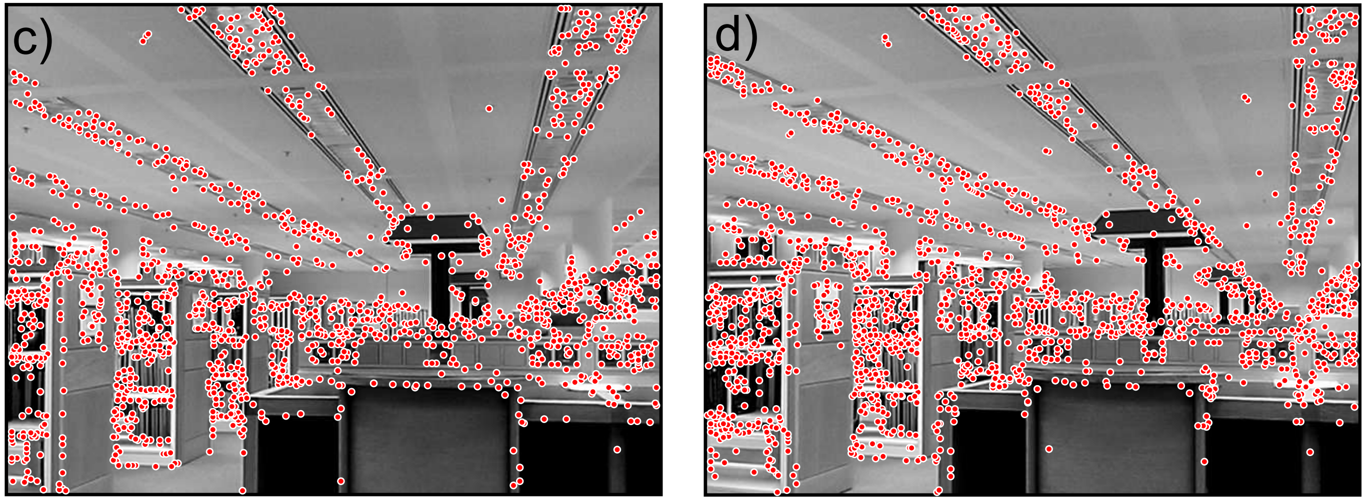
Source: Hartley and Zisserman, 2004

Summary

- Introduction
- 3D shapes from 2D images
- Stereo vision
- Correspondence
- Epipolar geometry
- The fundamental matrix
- The essential matrix
- RANSAC
- 3D reconstruction

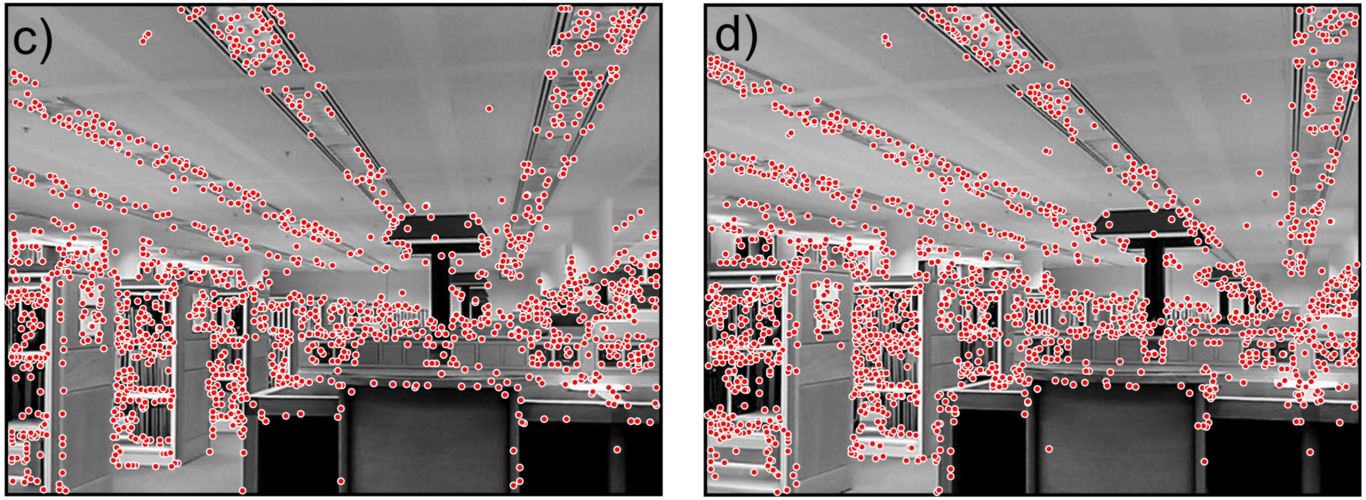
Two-view reconstruction pipeline

1. Compute image features
    - Find salient points in each image an interest point detector such as the SIFT detector



Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
   - Characterise the region around each feature in each image with a low dimensional vector, e.g. use the SIFT descriptor

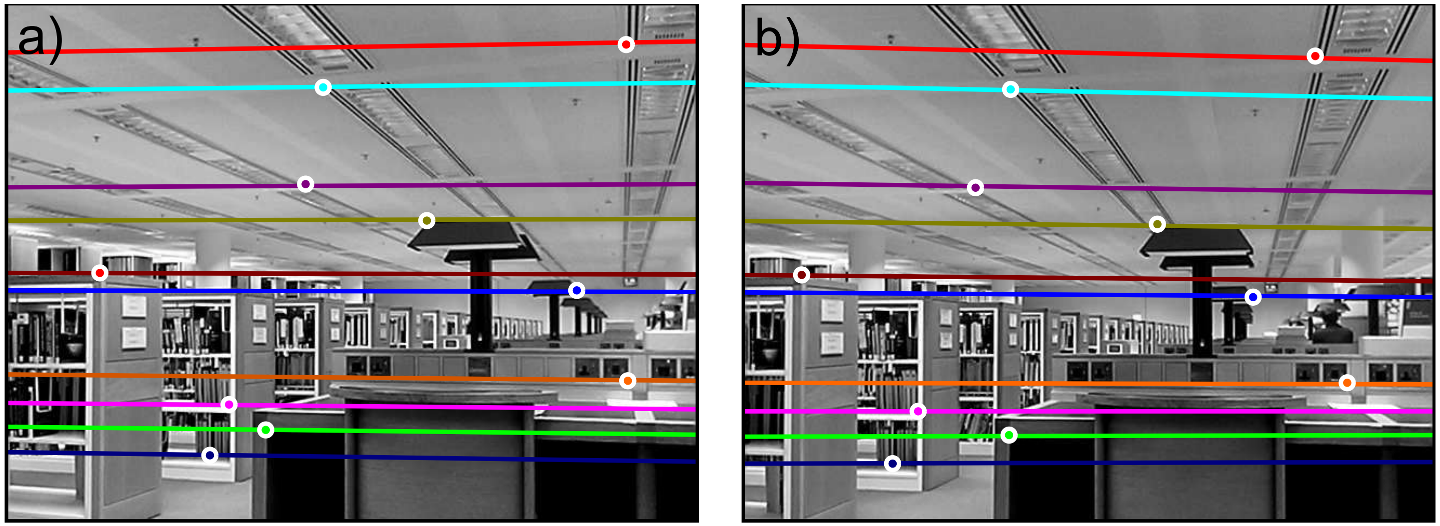


Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
   - Greedly match features between two features
   - E.g. Compute the squared distance between their region descriptors and stop when squared distance exceeds a pre-defined threshold.

Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
   - Compute the fundamental matrix using the eight-point algorithm or RANSAC

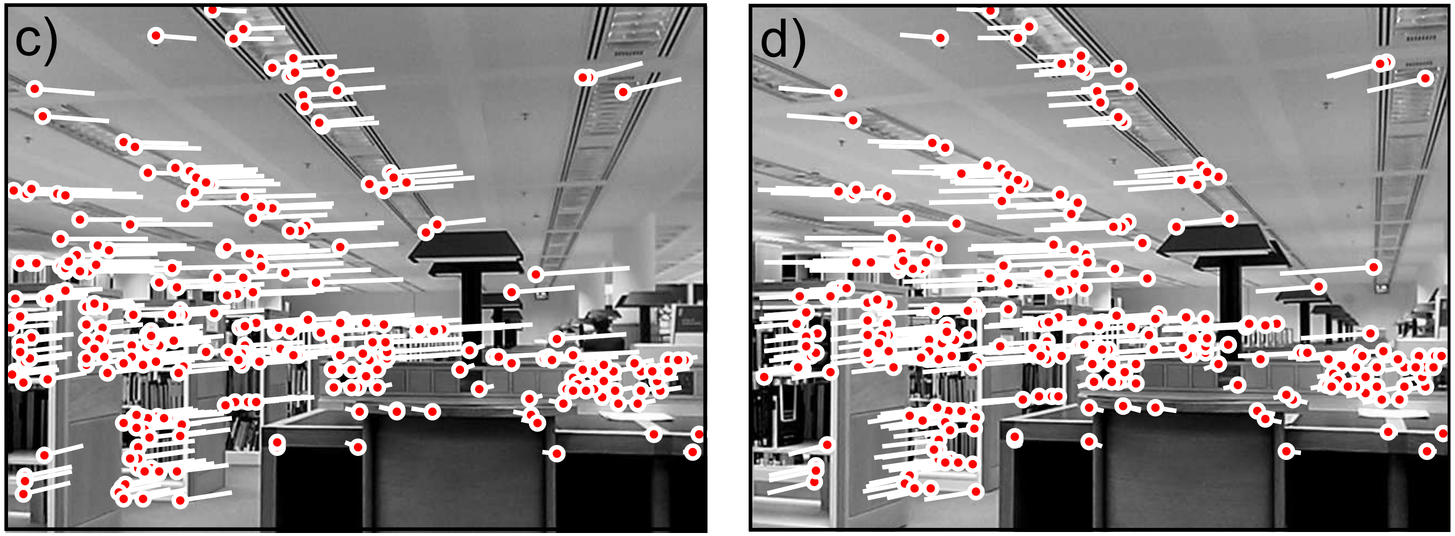


Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
5. Refine matches
    - Again greedy match features, but using epipolar geometry: If a putative match is not close to the induced epipolar line, it is rejected.
    - Recompute the fundamental matrix based on all the remaining point matches

Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
5. Refine matches



Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
5. Refine matches
6. Estimate essential matrix
    - Use intrinsic matrices from both cameras

$$\mathbf{E} = \mathbf{M_{int2} F M_{int1}}$$

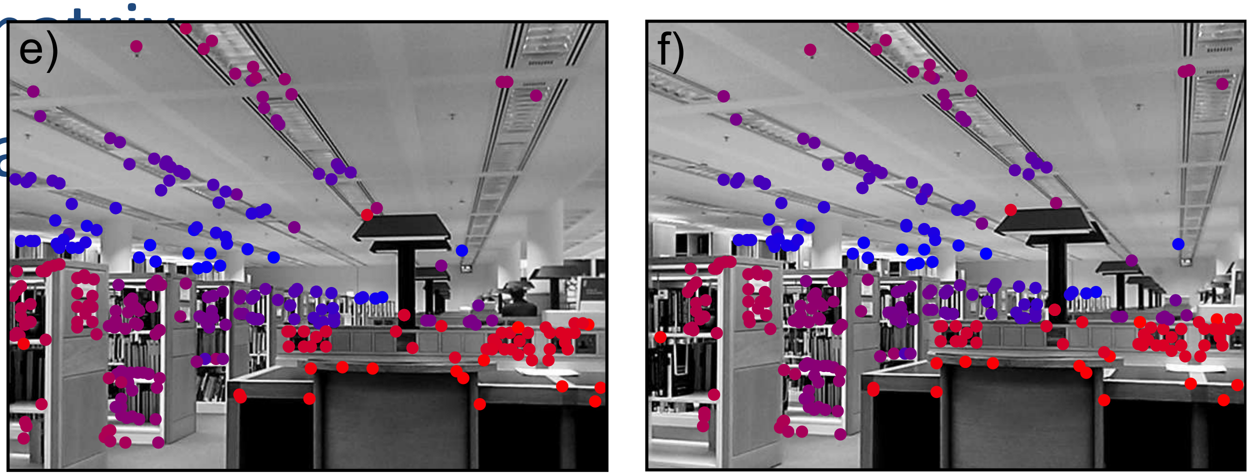
Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
5. Refine matches
6. Estimate essential matrix
7. Decompose essential matrix

$$\mathbf{E} = \tau_\times \mathbf{\Omega} \qquad \begin{aligned} \tau_\times &= \mathbf{ULWU}^T \\ \mathbf{\Omega} &= \mathbf{UW}^{-1}\mathbf{V}^T \end{aligned}$$

Two-view reconstruction pipeline

1. Compute image features
2. Compute feature descriptors
3. Find initial matches
4. Compute fundamental matrix
5. Refine matches
6. Estimate essential m
7. Decompose essenti
8. Estimate 3D points



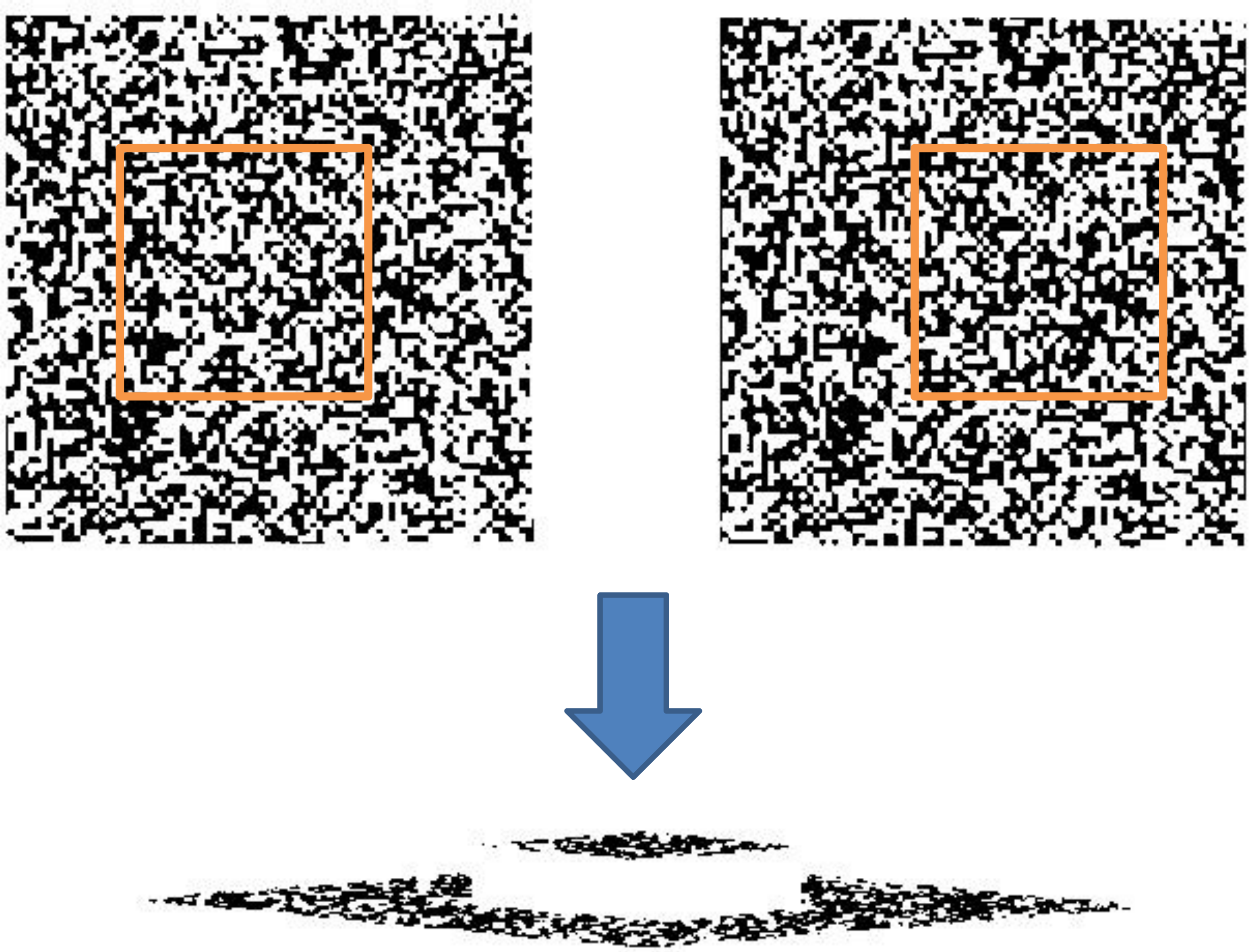
3D reconstruction

- Goal
  - Given binocular stereo images and their correspondences.
  - Recover the 3D geometry of the imaged scene.
- What can be achieved
  - Given intrinsic and extrinsic geometry of the cameras: Absolute Euclidean reconstruction.
  - Given only intrinsic camera geometries: Reconstruction up to a scale factor.
  - Given no information on camera geometries: Reconstruction up to a projective transformation.

Random-dot stereogram

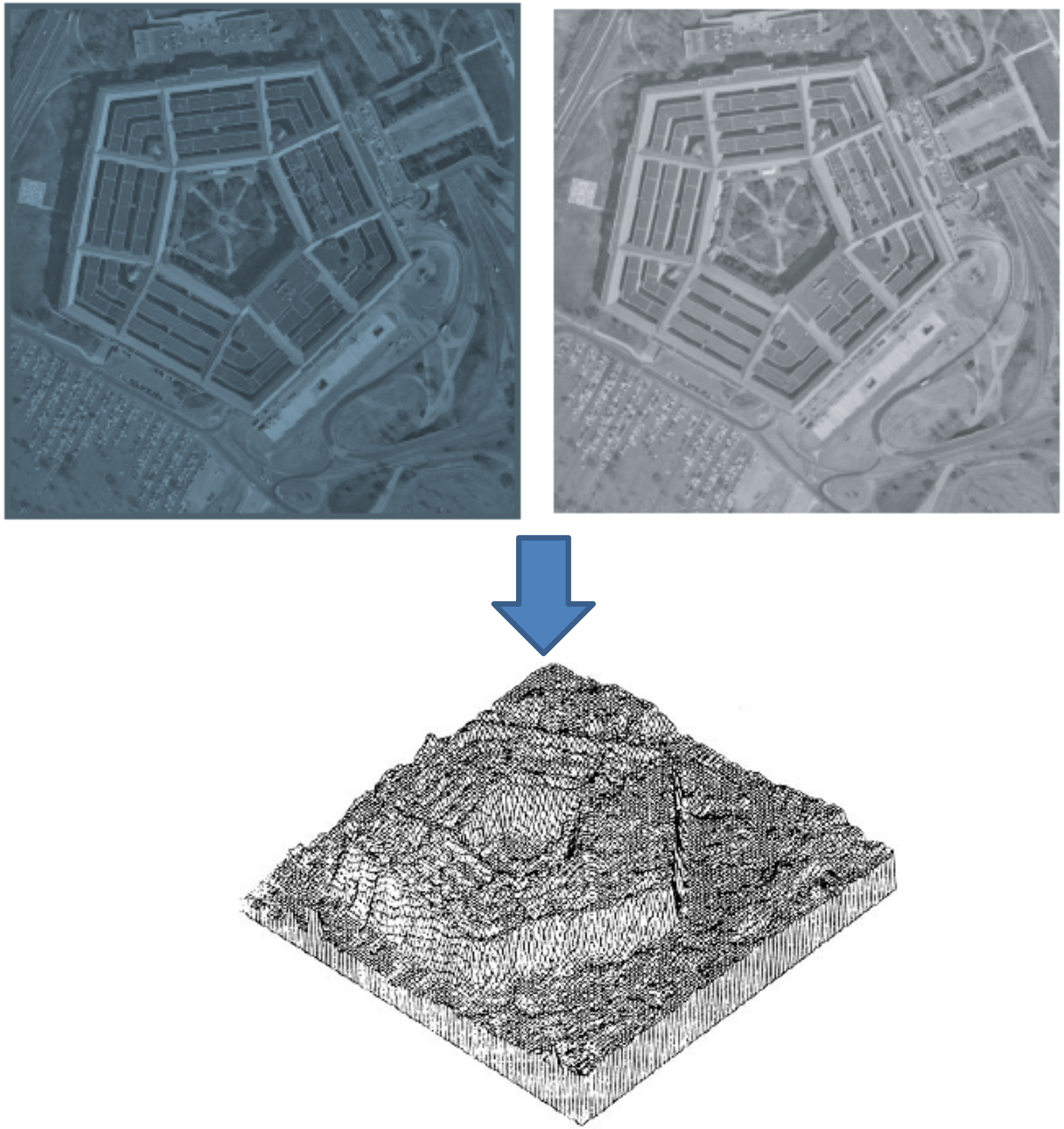
Random-dot stereogram

Random-dot stereogram



Real-world scene

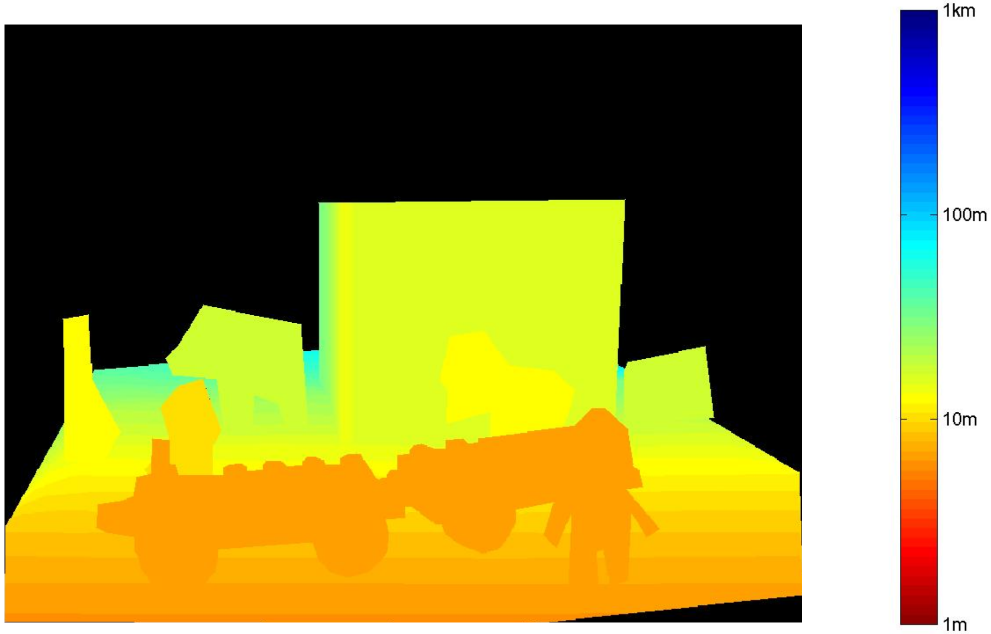
Real-world scene



Laboratory-calibrated scene

Source: Antonio Torralba slides

Laboratory-calibrated scene

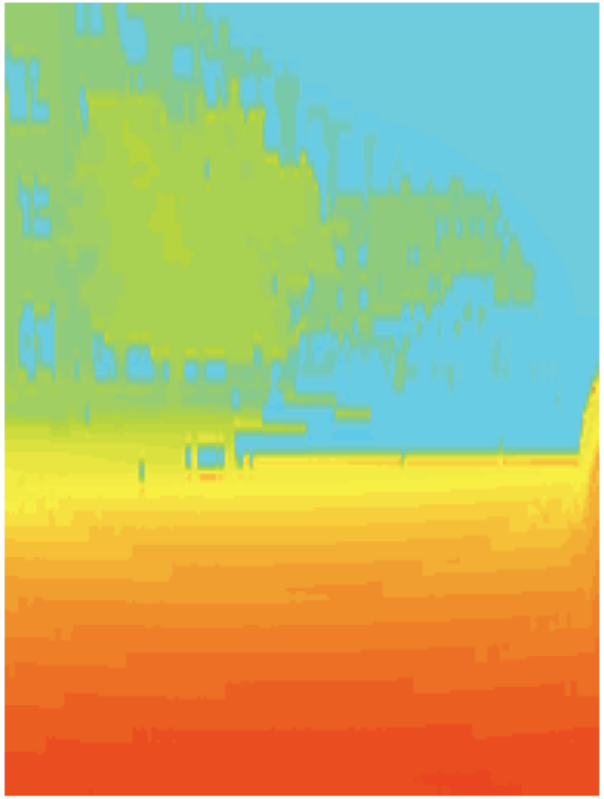


Source: Antonio Torralba slides

Street scene I

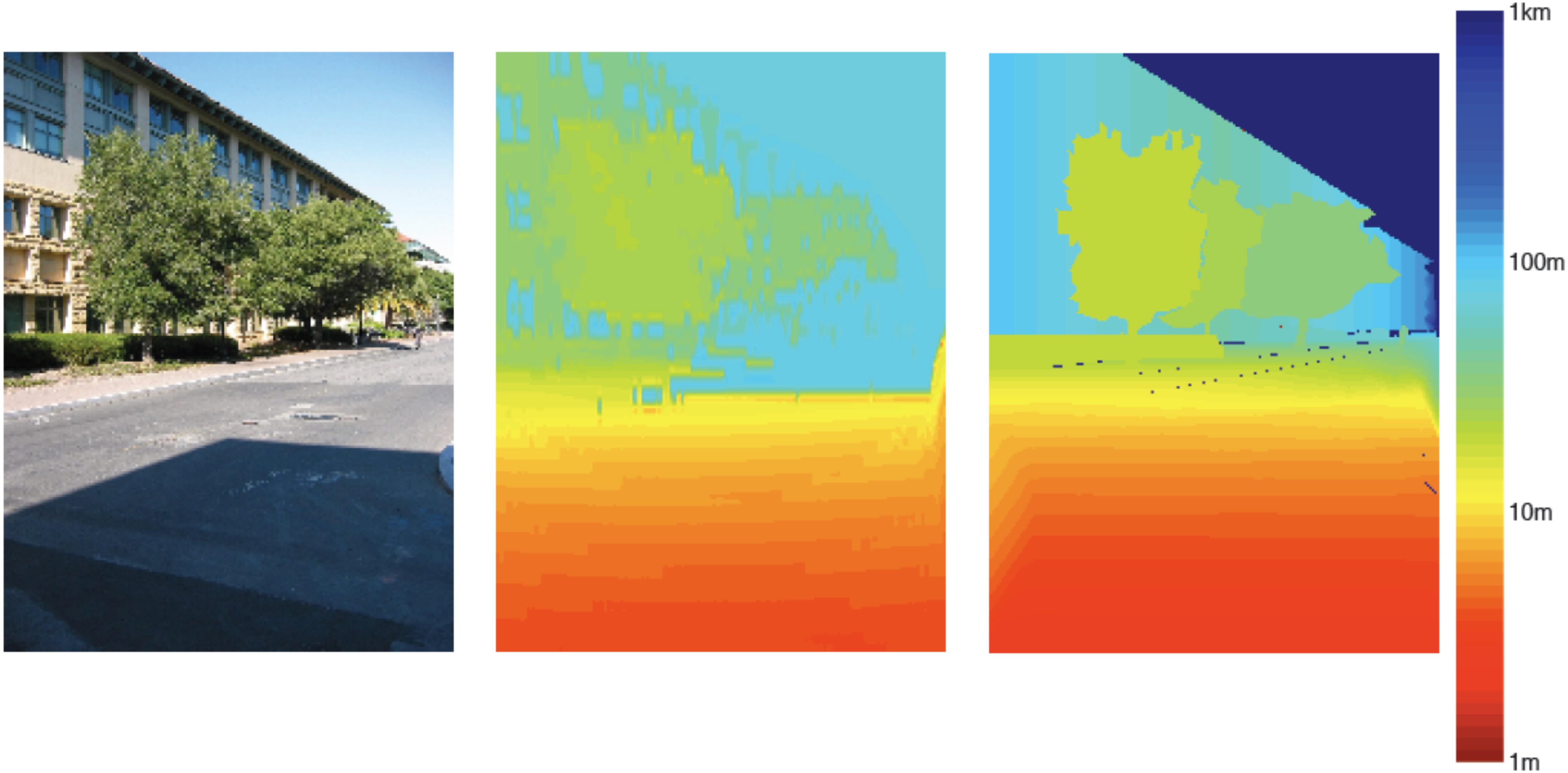
Source: Antonio Torralba slides

Street scene I



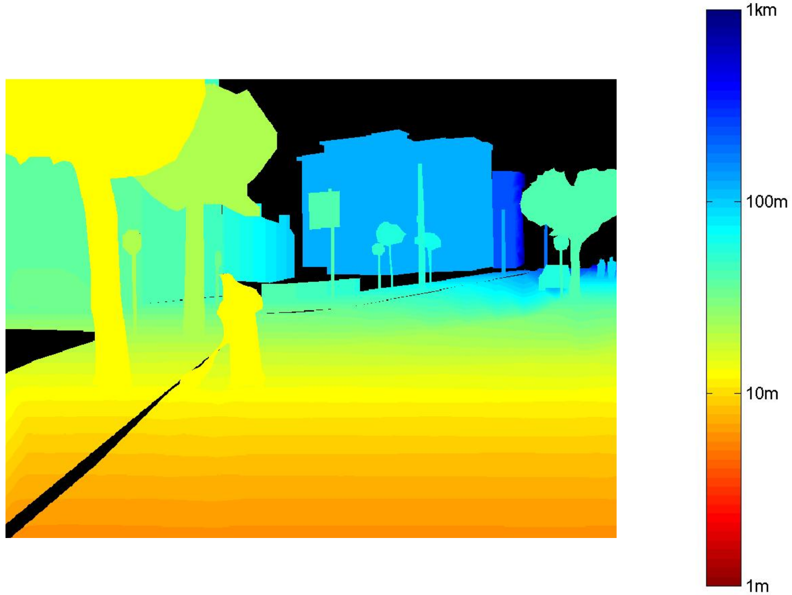
Source: Antonio Torralba slides

Street scene I



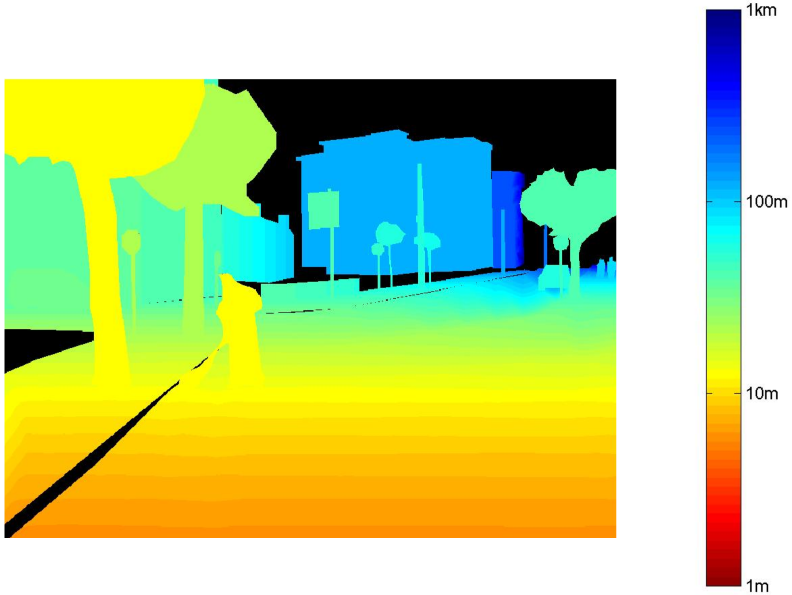
Source: Antonio Torralba slides

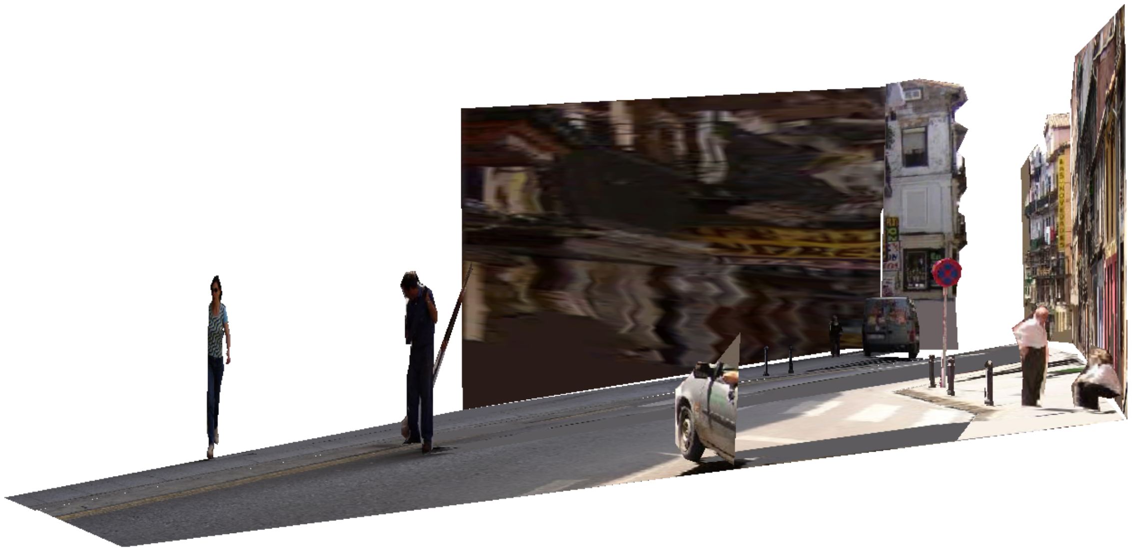
Street scene II



Source: Antonio Torralba slides

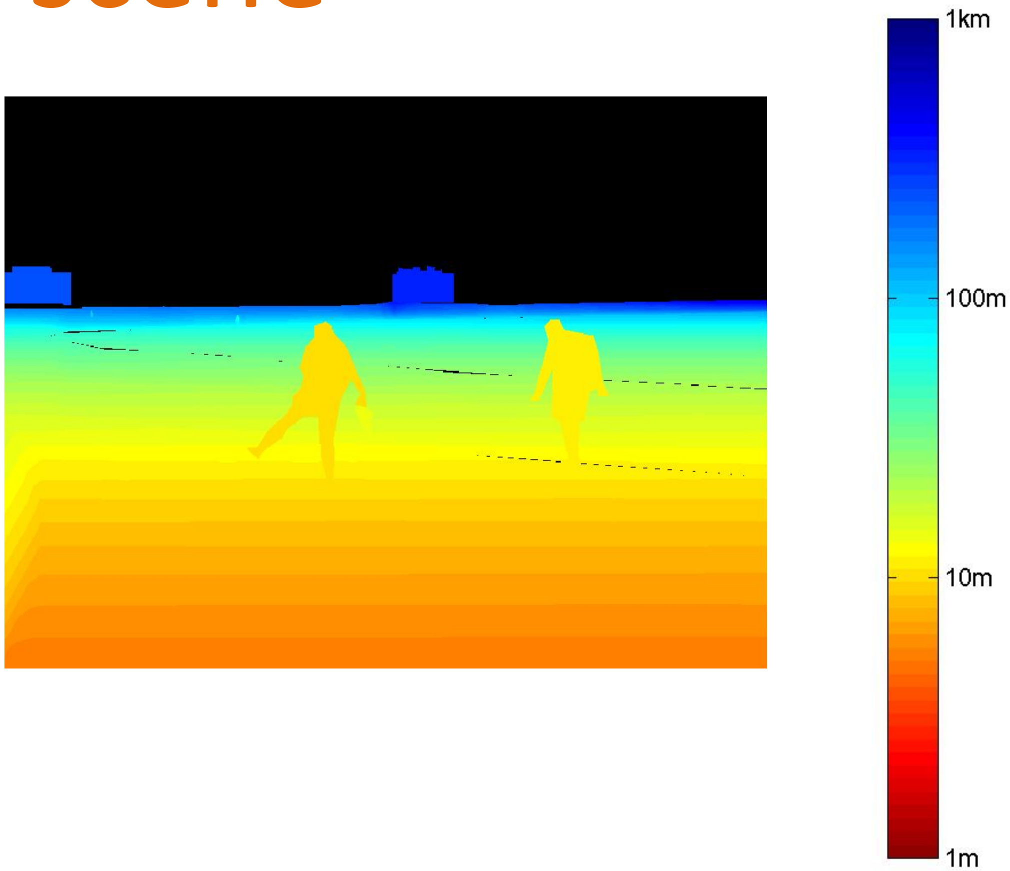
Street scene II





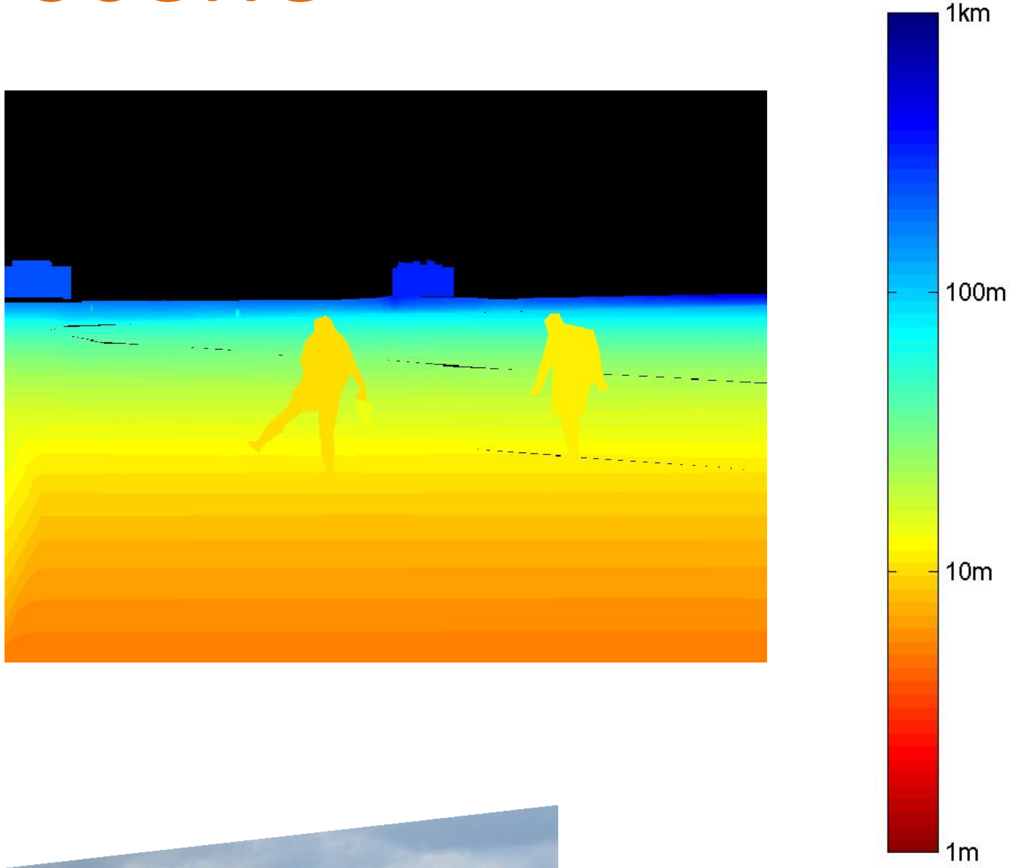
Source: Antonio Torralba slides

Beach scene



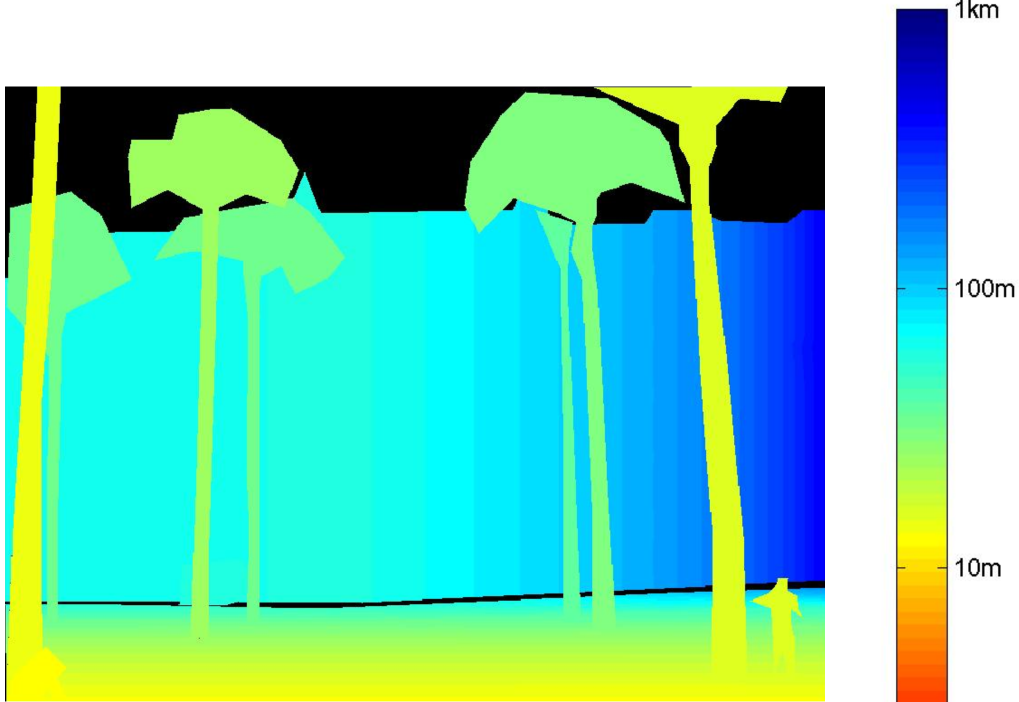
Source: Antonio Torralba slides

Beach scene



Source: Antonio Torralba slides

Building scene



Source: Antonio Torralba slides

Example: Multiview stereo

Source: Antonio Torralba slides

Summary

- Introduction
- 3D shapes from 2D images
- Stereo vision
- Correspondence
- Epipolar geometry
- The fundamental matrix
- The essential matrix
- RANSAC
- 3D reconstruction