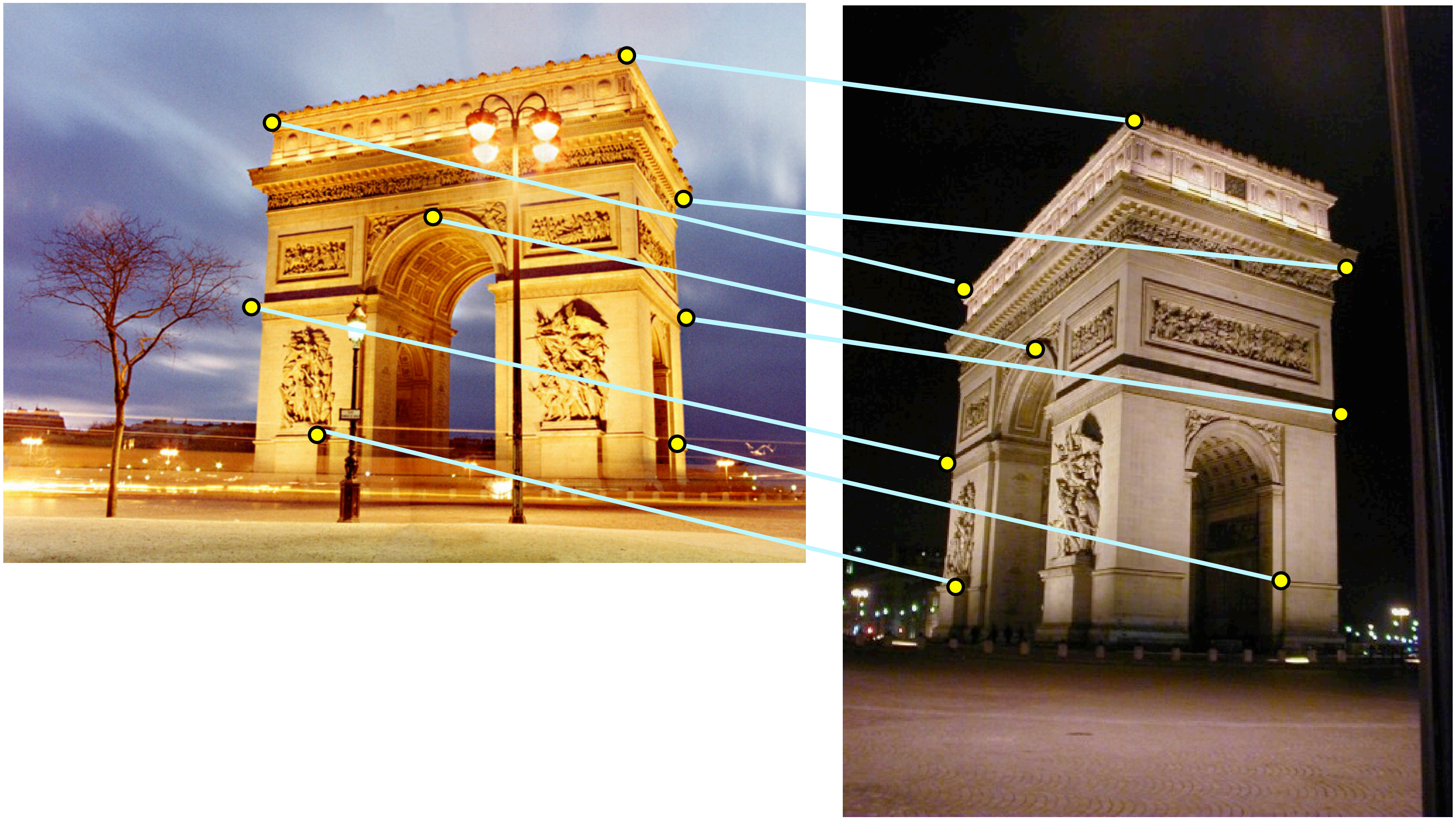
# Lecture 21:

# Stereo Reconstruction

Robert Collins
CSE486, Penn State

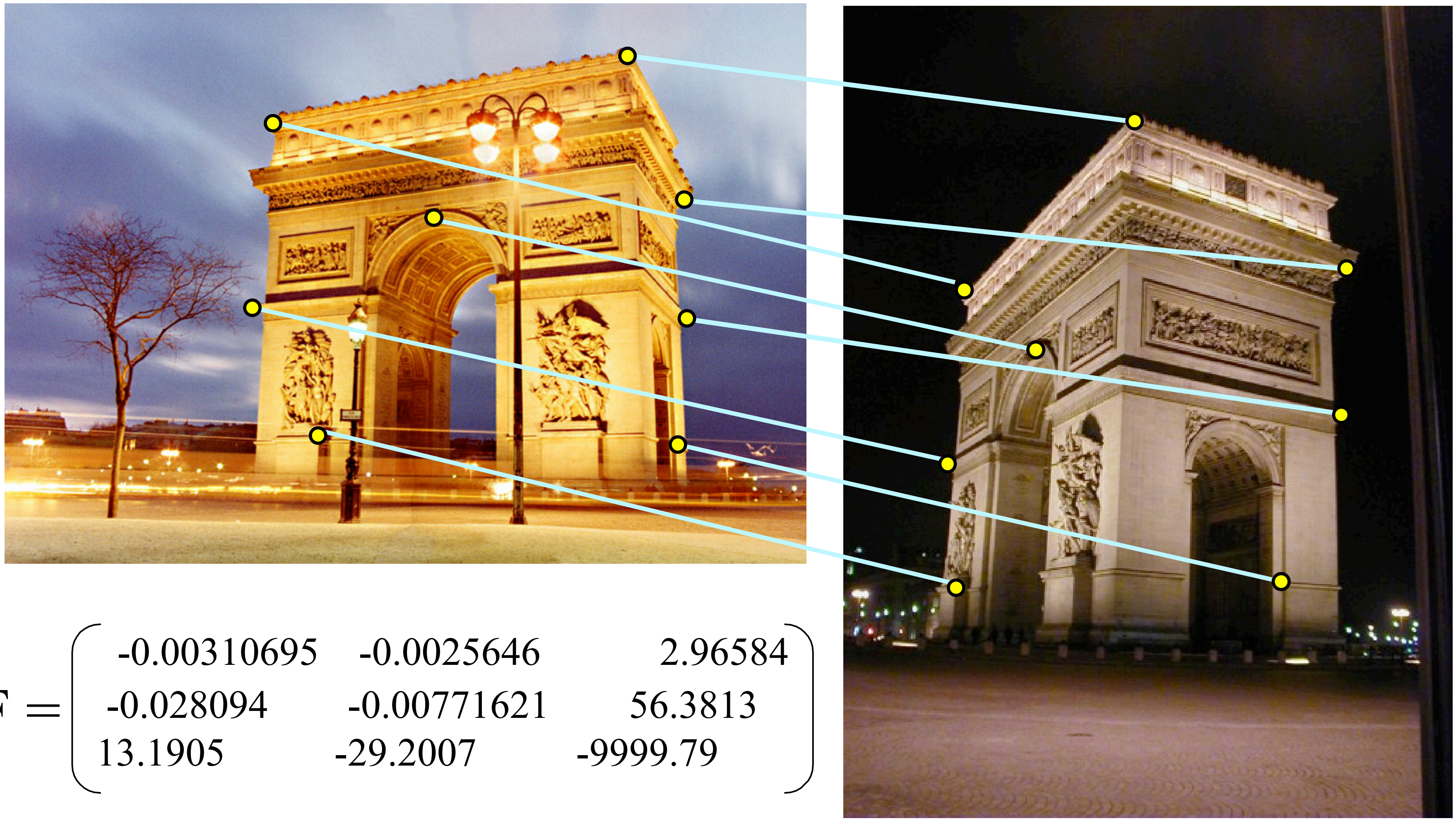
# Steps to General Stereo

find 8 or more initial point matches (somehow)



# Steps to General Stereo

## Compute F matrix using 8-point algorithm

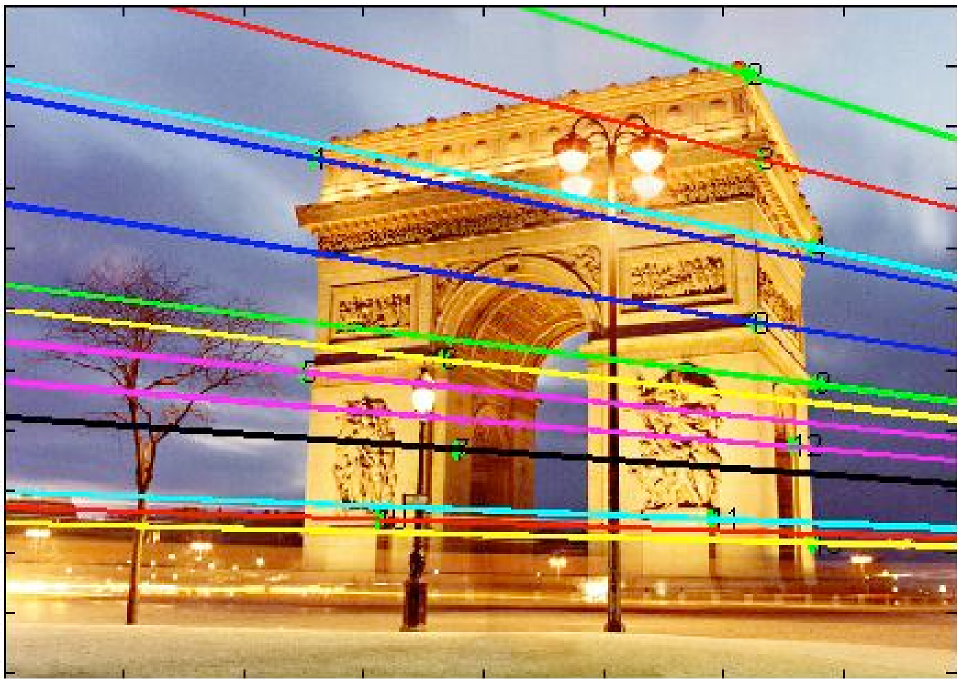


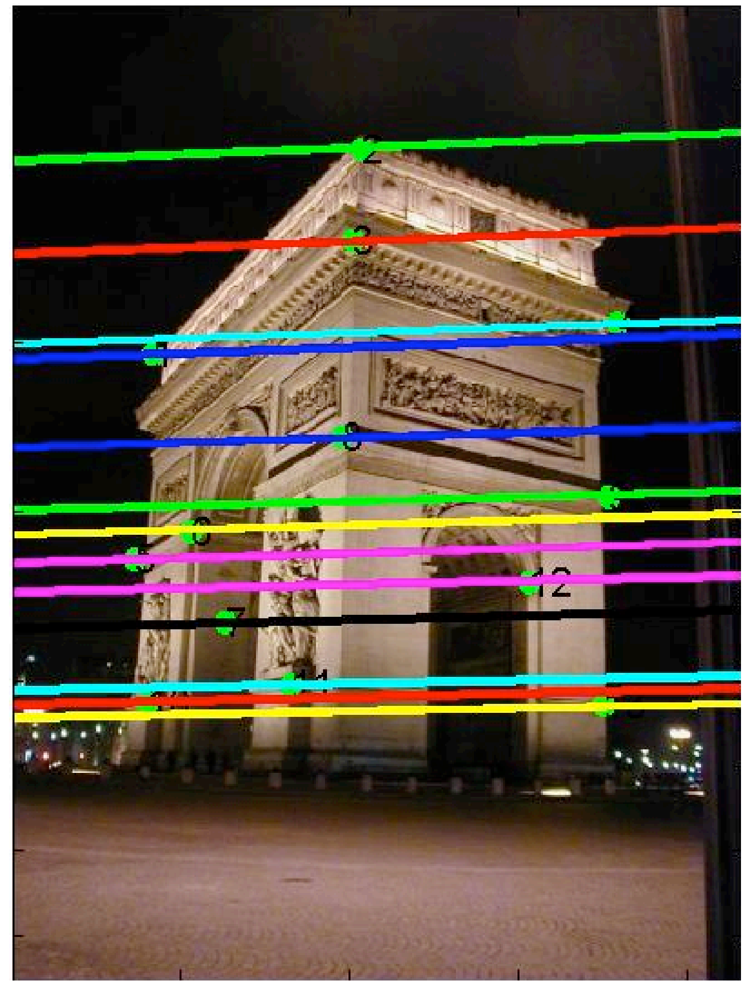
$$F = \begin{pmatrix} -0.00310695 & -0.0025646 & 2.96584 \\ -0.028094 & -0.00771621 & 56.3813 \\ 13.1905 & -29.2007 & -9999.79 \end{pmatrix}$$

Robert Collins
CSE486, Penn State

# Steps to General Stereo

Infer epipolar geometry (epipoles, epipolar lines) from F.

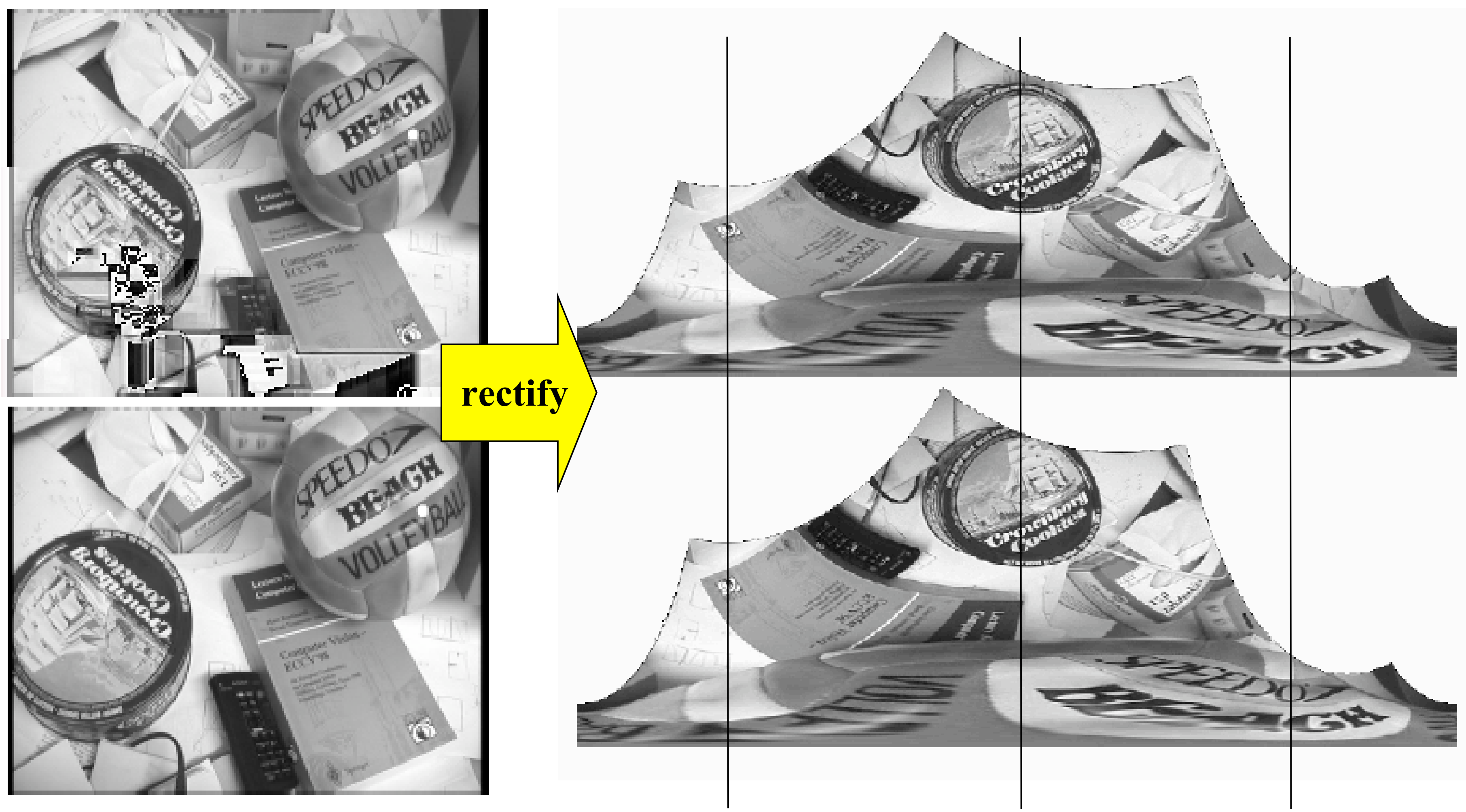




$$F = \begin{pmatrix} -0.00310695 & -0.0025646 & 2.96584 \\ -0.028094 & -0.00771621 & 56.3813 \\ 13.1905 & -29.2007 & -9999.79 \end{pmatrix}$$

# Steps to General Stereo

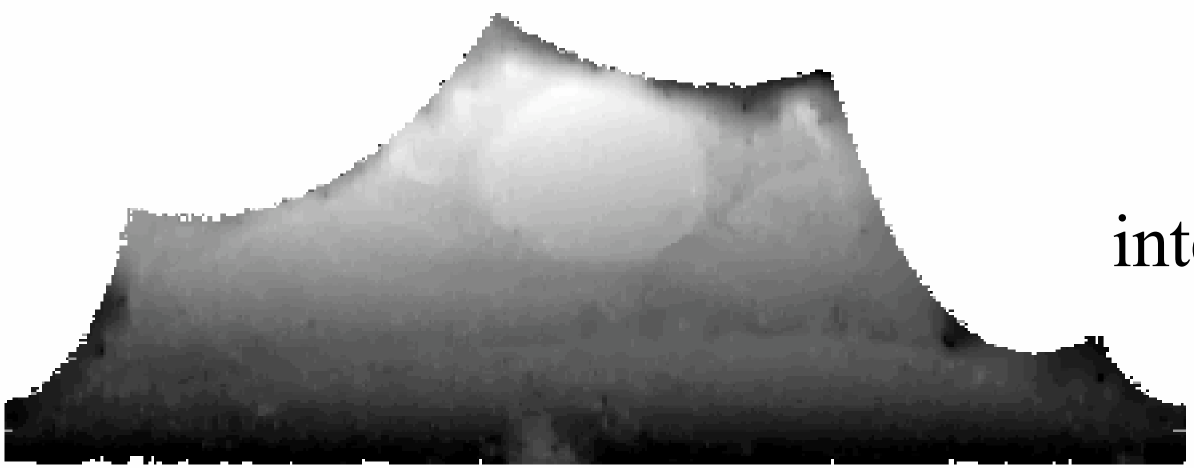
Rectify images to get simple scanline stereo pair.



# Steps to General Stereo

Compute disparity map (correspondence matching)

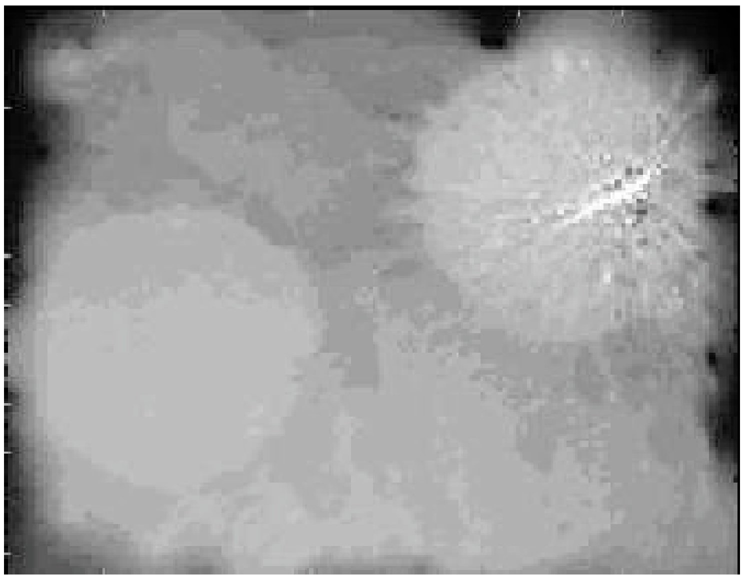
sparse



interpolated

# Steps to General Stereo

Compute 3D surface geometry from disparity map

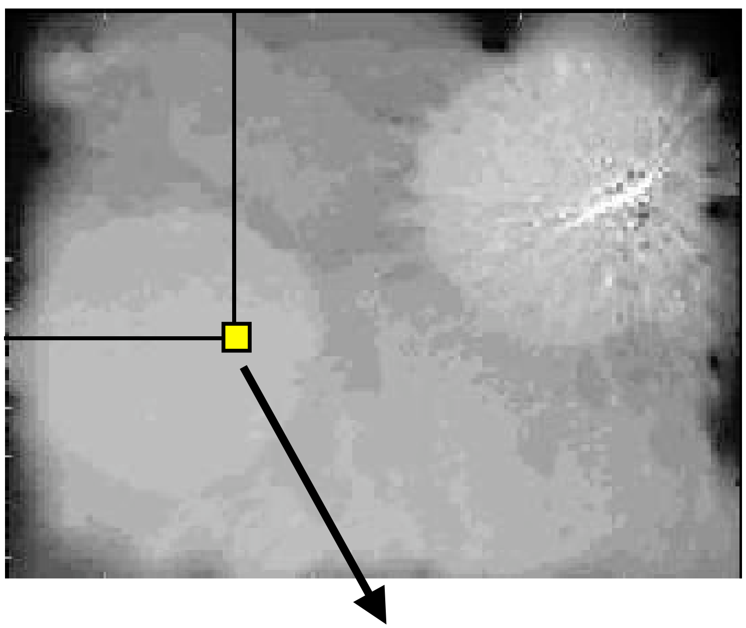


disparity map in
pixel coords


Views of texture mapped
depth surface

# Reconstruction

Even though we have a dense disparity map, when talking about recovering 3D scene structure, we will consider it to just be a set of point matches.




<u>point match:</u>

(u,v) in left image
matches
(u-d,v) in right image

# Stereo Reconstruction

Given point correspondences, how to compute 3D point positions using triangulation.

<u>Results depend on how calibrated the system is:</u>

1) Intrinsic and extrinsic parameters known
   Can compute metric 3D geometry
2) Only intrinsic parameters known
   Unknown scale factor
3) Neither intrinsic nor extrinsic known
   Recover structure up to an unknown projective transformation of the scene

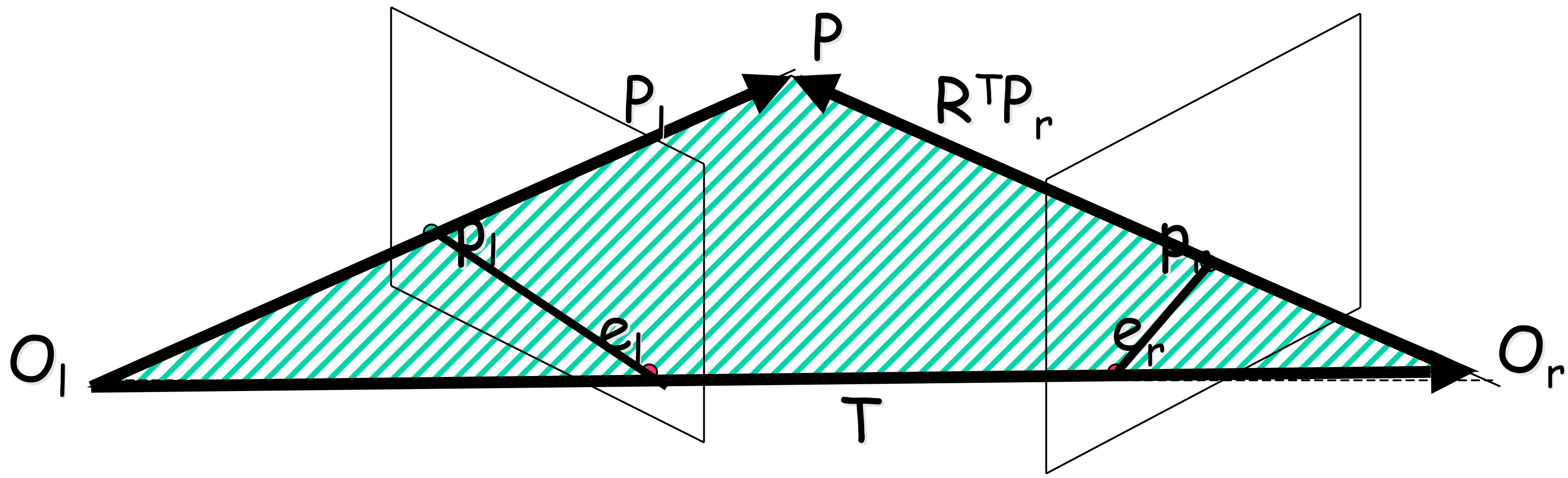
# Fully Calibrated Stereo

Known intrinsics -- can compute viewing rays in camera coordinate system

Know extrinsics -- know how rays from both cameras are positioned in 3D space

Reconstruction: triangulation of viewing rays

# Calibrated Triangulation



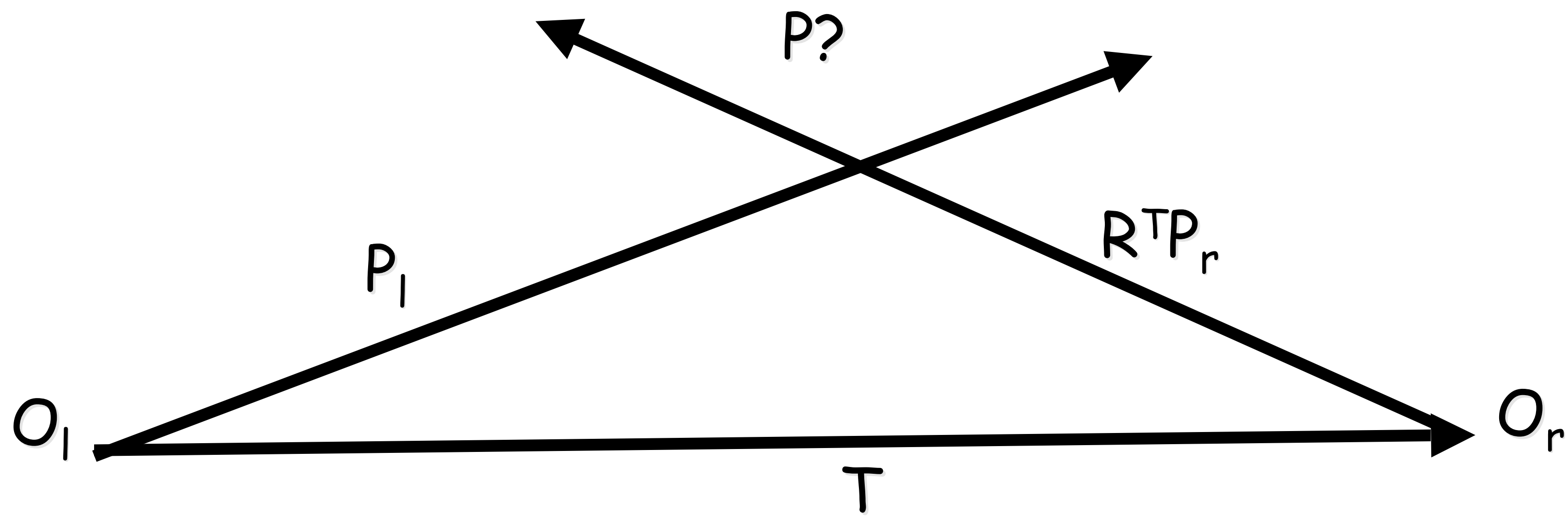

ideally, P is the point of intersection of two 3D rays:

ray through $O_l$ with direction $P_l$

ray through $O_r$ with direction $R^TP_r$

# Triangulation with Noise

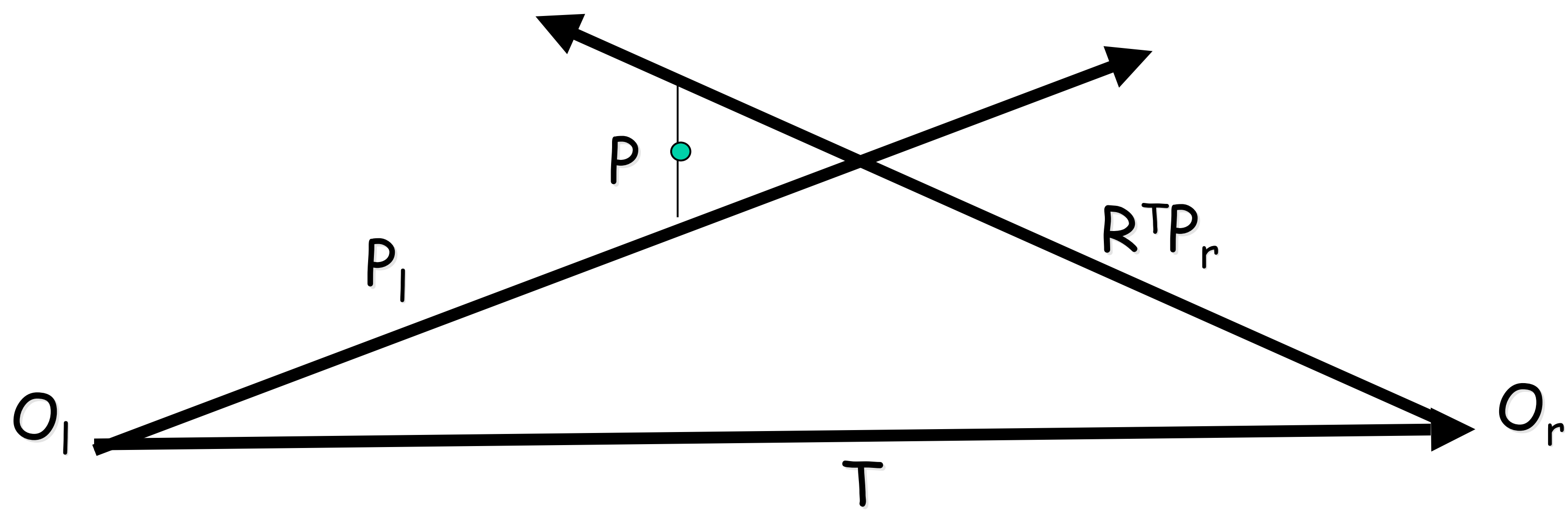
Unfortunately, these rays typically don't intersect due to noise in point locations and calibration params



Robert Collins
CSE486, Penn State

# Triangulation with Noise

Unfortunately, these rays typically don't intersect due to noise in point locations and calibration params

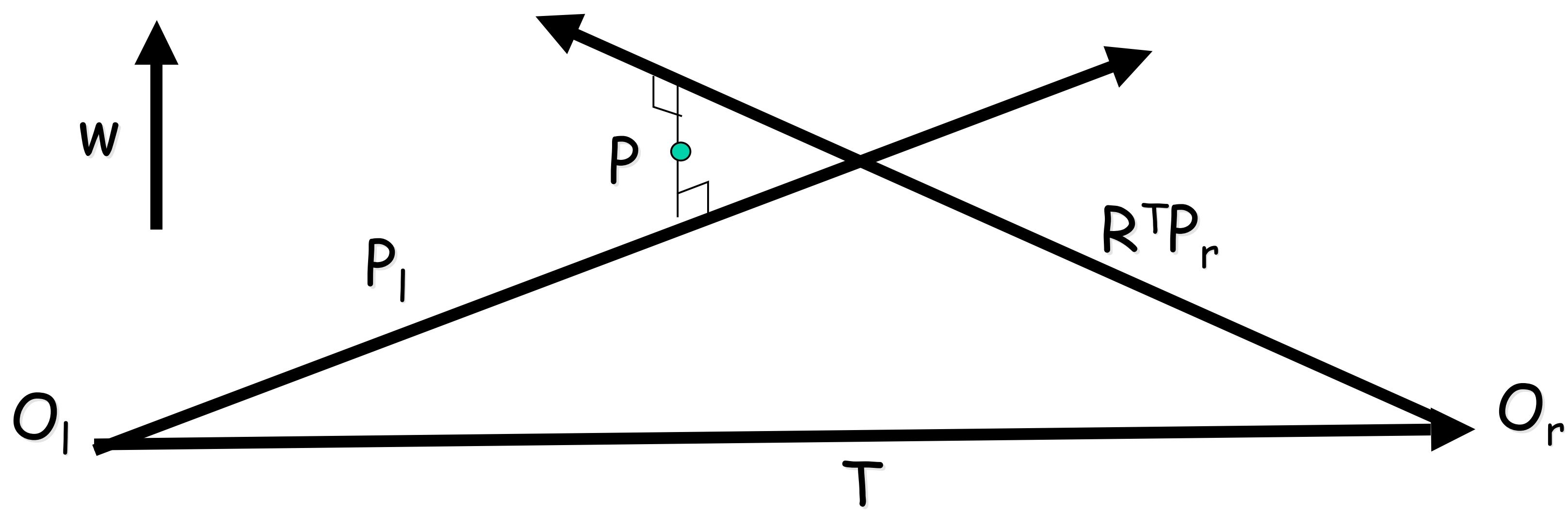



Solution: Choose P as the "pseudo-intersection point". This is point that minimizes the sum of squared distance to both rays. (The SSD is 0 if the rays exactly intersect)

# Solution from T&V Book

P is midpoint of the segment perpendicular to $P_l$ and $R^T P_r$

Let $w = P_l \times R^T P_r$ (this is perpendicular to both)



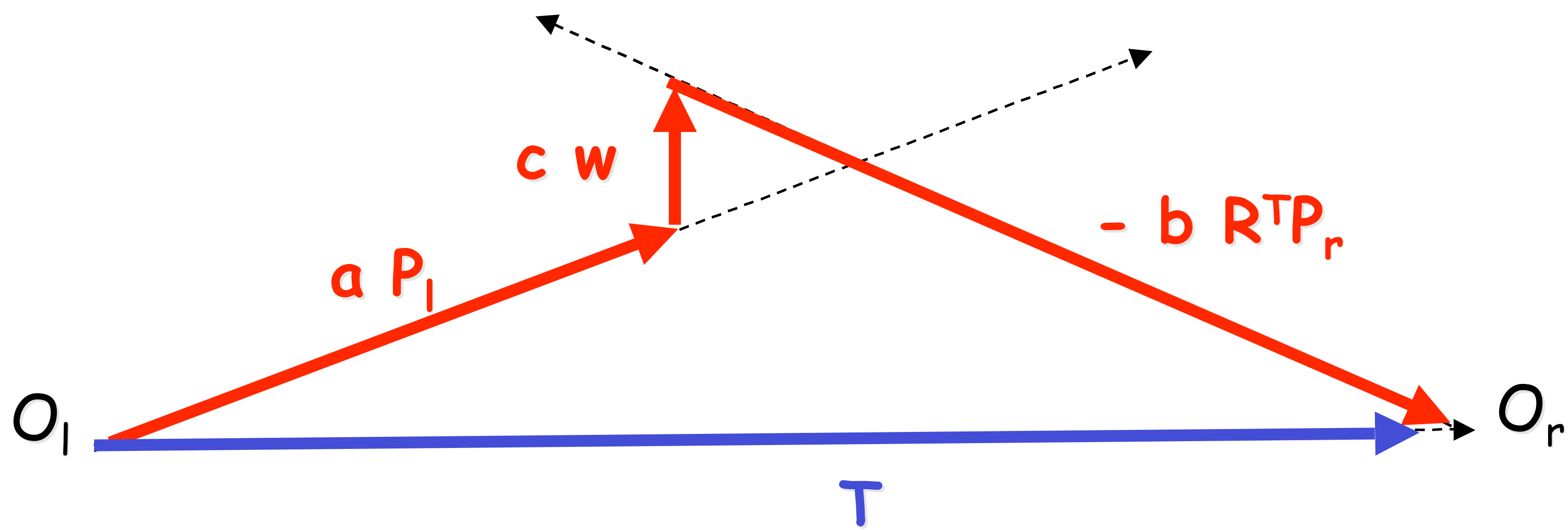

Introducing three unknown scale factors a,b,c we note we can write down the equation of a "circuit"

Robert Collins
CSE486, Penn State

# Solution from T&V Book

Writing vector "circuit diagram" with unknowns a,b,c

$$a\,P_l + c\,(P_l \times R^T P_r) - b\,R^T P_r = T$$

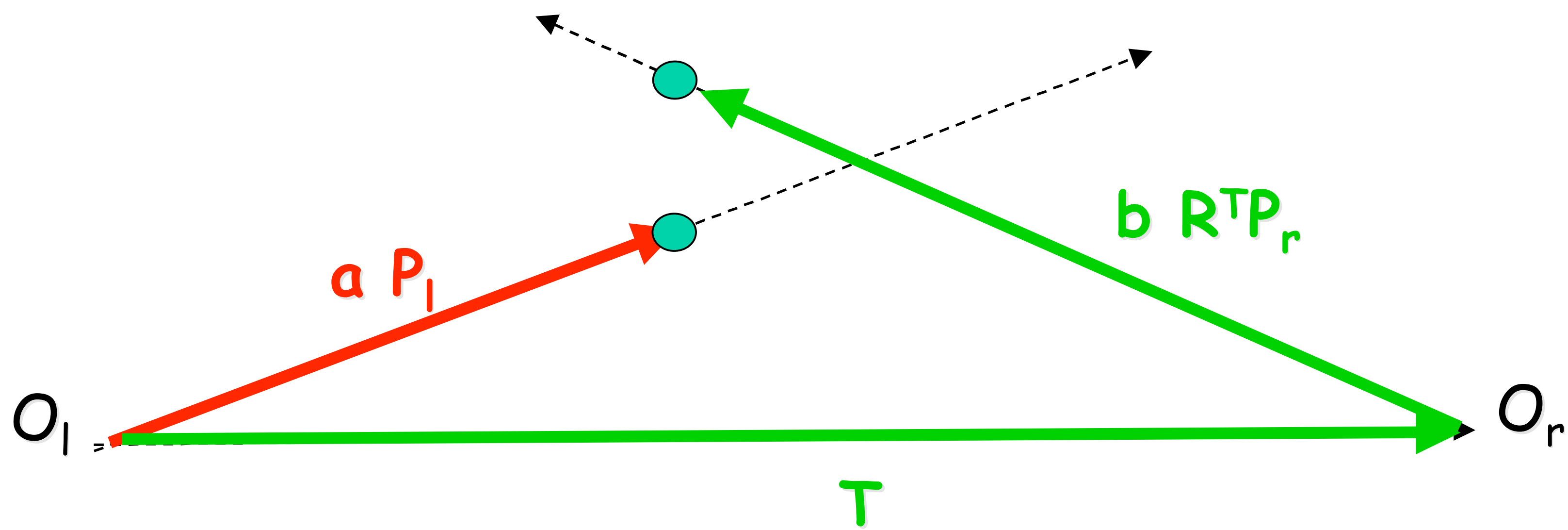



note: this is three linear equations in three unknowns a,b,c
=> can solve for a,b,c

Robert Collins
CSE486, Penn State

# Solution from T&V Book

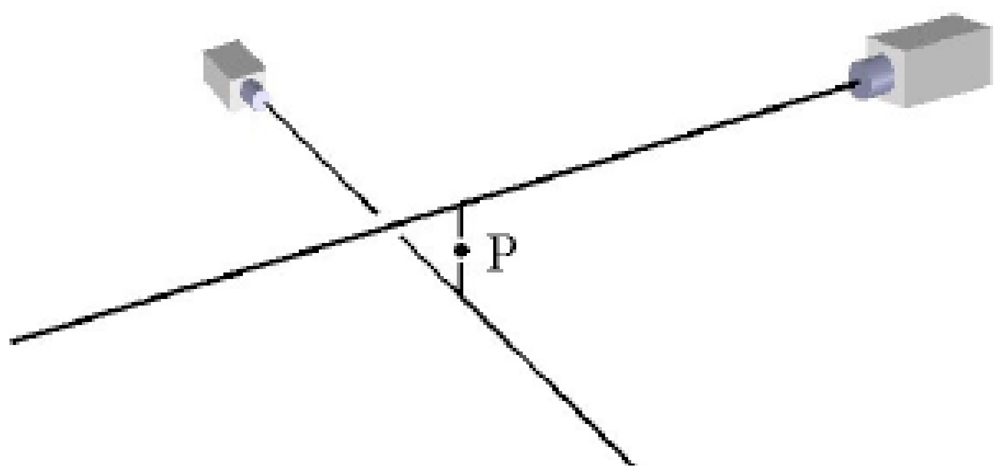
After finding a,b,c , solve for midpoint of line segment between points $O_l + a\,P_l$ and $O_l + T + b\,R^T P_r$



Robert Collins
CSE486, Penn State

# Alternate Solution

I prefer an alternate solution based on using least squares to solve for an unknown point P that minimizes SSD to viewing rays. Why? it generalizes readily to N cameras.




$$[\sum_{i}^{n} w_i(I - u_i u_i')]\, P \;=\; \sum_{i}^{n} w_i(I - u_i u_i')c_i$$

# Stereo Reconstruction

Given point correspondences, how to compute 3D point positions using triangulation.

Results depend on how calibrated the system is:

1) Intrinsic and extrinsic parameters known
   Can compute metric 3D geometry
2) Only intrinsic parameters known
   Unknown scale factor
3) Neither intrinsic nor extrinsic known
   Recover structure up to an unknown projective transformation of the scene

# Only Intrinsic Params Known

General outline of solution (see book for details)

Use knowledge that E = R S to solve for R and T, then use previous triangulation method.

Note: since E is only defined up to a scale factor, we can only determine the direction of T, not its length. So... 3D reconstruction will have an unknown scale.

# Only Intrinsic Params Known

Using E to solve for extrinsic params R and T

$E = R\ S$ where elements of S are functions of T

Then $E^T E = S^T R^T R\ S = S^T S$ (because $R^T R = I$)

Thus $E^T E$ is only a function of T.
Solve for elements of T assuming it is a unit vector.

After determining T, plug back into $E = R\ S$ to determine R.

# Only Intrinsic Params Known

Unfortunately, four different solutions for (R,T) are possible (due to choice of sign of E, and choice of sign of T when solving for it).

However, only one choice will give consistent solutions when used for triangulation, where <u>consistent</u> means reconstructed points are in front of the cameras (positive Z coordinates).

So, check all four solutions, choose the correct one, and you are done.

# Stereo Reconstruction

Given point correspondences, how to compute 3D point positions using triangulation.

Results depend on how calibrated the system is:

1) Intrinsic and extrinsic parameters known
   Can compute metric 3D geometry
2) Only intrinsic parameters known
   Unknown scale factor
3) Neither intrinsic nor extrinsic known
   Recover structure up to an unknown projective transformation of the scene

# Stereo when "Nothing" is Known

What if we don't known intrinsic nor extrinsic params? (we just look at two pictures of the scene with no prior information)

Can we recover any 3D information?

It wasn't clear that you could, but then in 1992...

- •Faugeras "What can be seen in three dimensions from an uncalibrated stereo rig", ECCV 1992
- •Hartley et.al., "Stereo from Uncalibrated Cameras", CVPR 1992
- •Mohr et.al., "Relative 3D Reconstruction using Multiple Uncalibrated Images, LIFIA technical report, 1992

# Stereo when "Nothing" is Known

Of course, we don't really know "nothing".

We know point correspondences, and because of that, we can compute the fundamental matrix F

Sketch of solution: use knowledge of F and use 5 points in the scene to define an arbitrary projective coordinate system. These points will have coordinates:

(1 0 0 0) (0 1 0 0) (0 0 1 0) (0 0 0 1) (1 1 1 1)

Result: You can recover 3D locations of other points with respect to that projective coordinate system.

# Stereo when "Nothing" is Known

Result: You can recover 3D locations of other points with respect to that projective coordinate system.

Why would that be practical?

Often, you can then use other information to determine how your arbitrary projective coordinate system relates to the "real" Euclidean scene coordinates system.
e.g. use prior knowledge of lengths and angles
of some items in the world (like a house)