# Fusion of Uncalibrated Sensor Streams

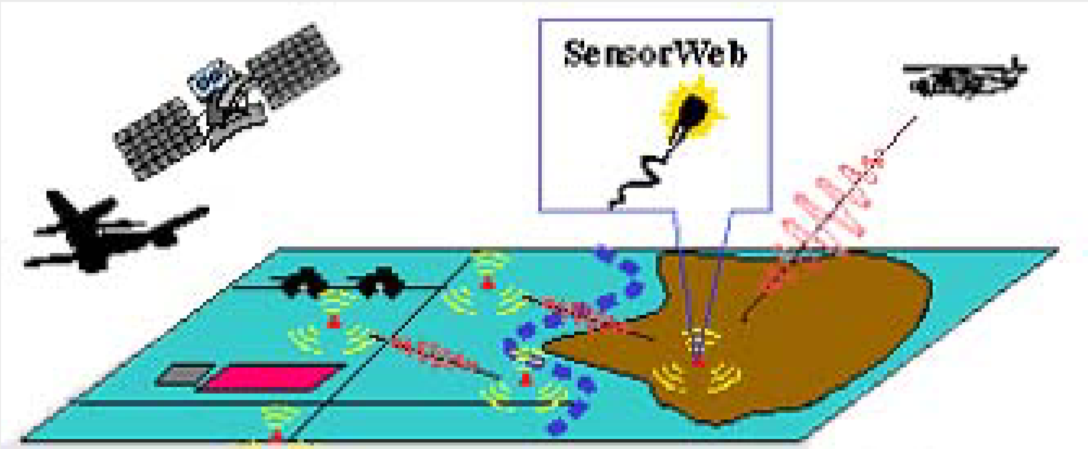
MURI Review Meeting
June 18, 2001

Sanjeev Kulkarni
Princeton University

# Outline

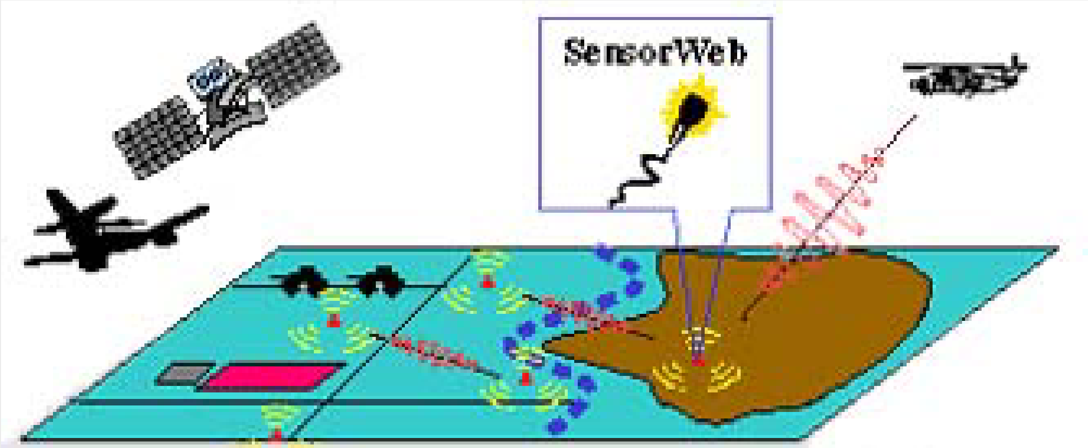
- Problem and motivation
- Estimating relative sensor geometry
- Obtaining detailed signal correspondences
- Recursive propagation and fusing dynamic streams
- Future directions

# Problem and Motivation




- Would like to fuse myopic information to attain more global view of battlefield scenario
- Dynamic scene/sensing: need fast algorithms but can exploit temporal regularities.
- Unknown scene and sensor geometry

# Problem and Motivation cont.

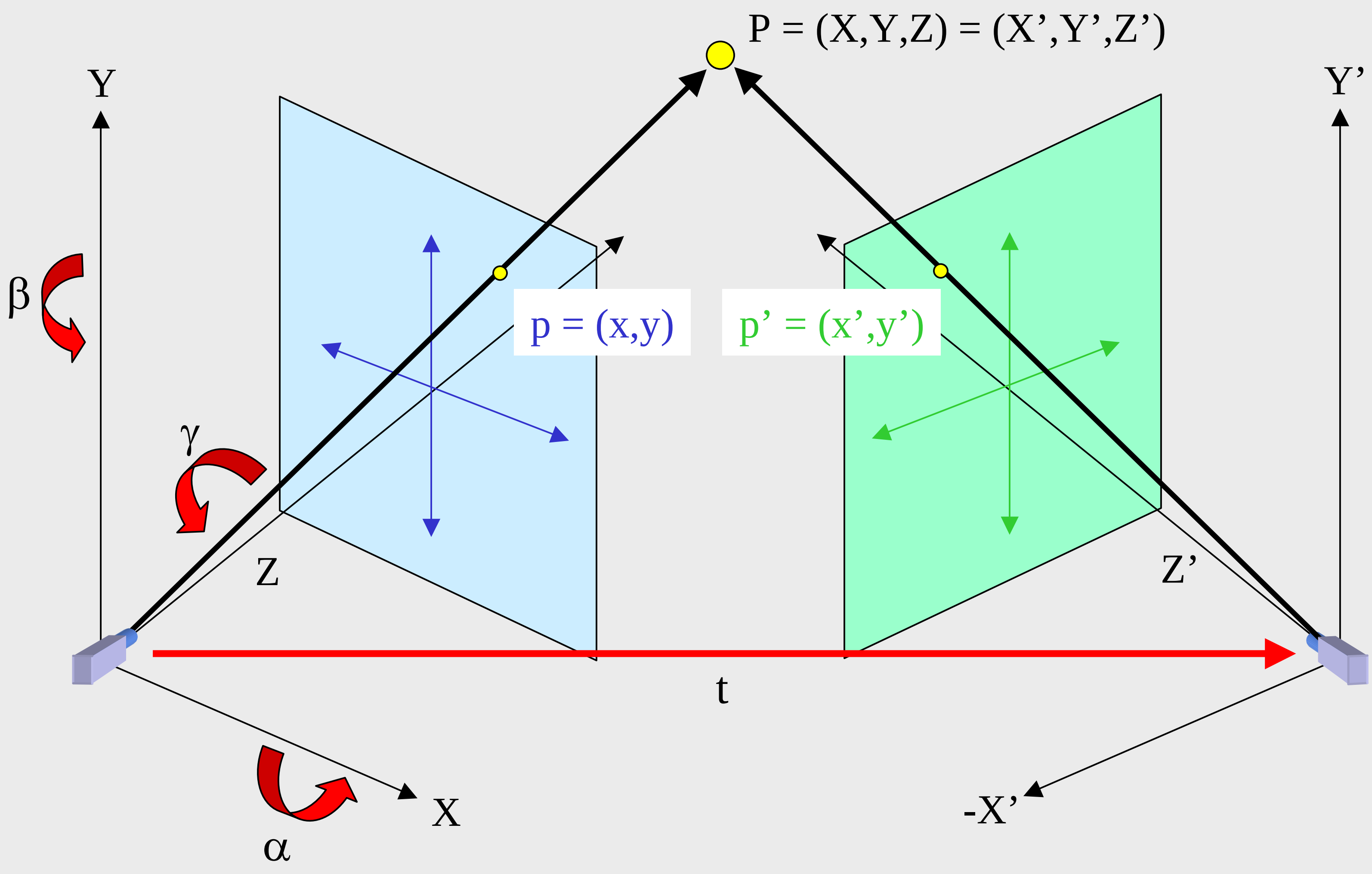



- Complex, dynamic environment
- Multiple, widely separated sensing
- Uncalibrated, possibly dynamic, sensors (unknown parameters, geometry)
- Noise

# Approach

- In-depth analysis with two sensor streams
- Use video as the sensing mode surrogate
- For a fixed snapshot, develop methods to estimate relative geometry.
- Exploit temporal regularities to develop fast recursive method to deal with dynamics

# Cameras and Images



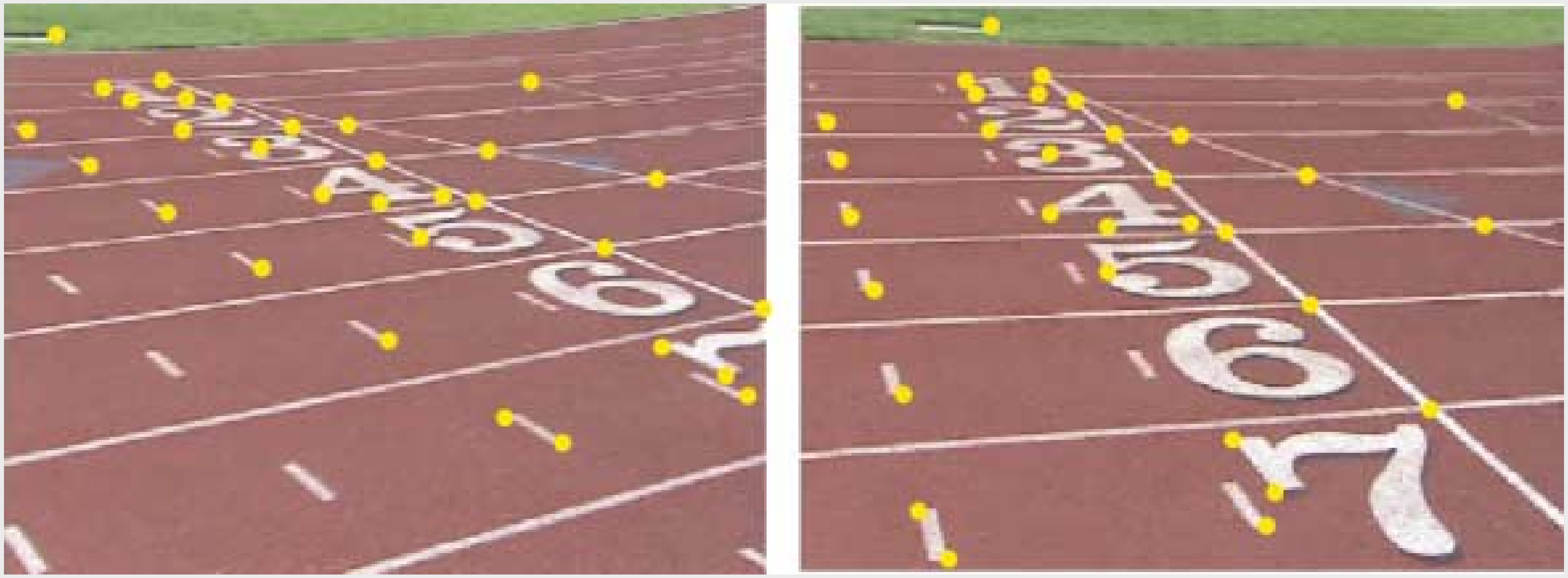
# Estimating Relative Sensor Geometry

- For camera rotation or planar scene, image points are related by a projective transformation:

$$g_M(w) = \frac{Aw + b}{c^T w + 1} \qquad A \in R^{2x2};\ \ b, c \in R^2$$

- Given two views of overlapping scene, would like to estimate the projective transformation.
- Typically an 8-dimensional non-quadratic minimization. We develop a 2-dimensional reduction that can be solved efficiently.

# Noisy Samples



Given a set of point mappings: $\left\{ w_j \mapsto w_j' \in R^2, \;\; j = 1, \ldots, N \right\}$

These are noisy samples of a fixed but unknown PT:

$$g_M^* : \;\; w_j' = g_M^*(w_j) + e_j, \;\; j = 1, \ldots, N$$

# The Least-Squares Estimate

$$\min_{M=\{A,b,c\}} Q(M) = \sum_{j=1}^{N} \left( w_j' - \frac{A\,w_j + b}{c^T w_j + 1} \right)^T \left( w_j' - \frac{A\,w_j + b}{c^T w_j + 1} \right)$$

Generally solved using a numerical minimization algorithm, e.g. Levenberg-Marquardt.

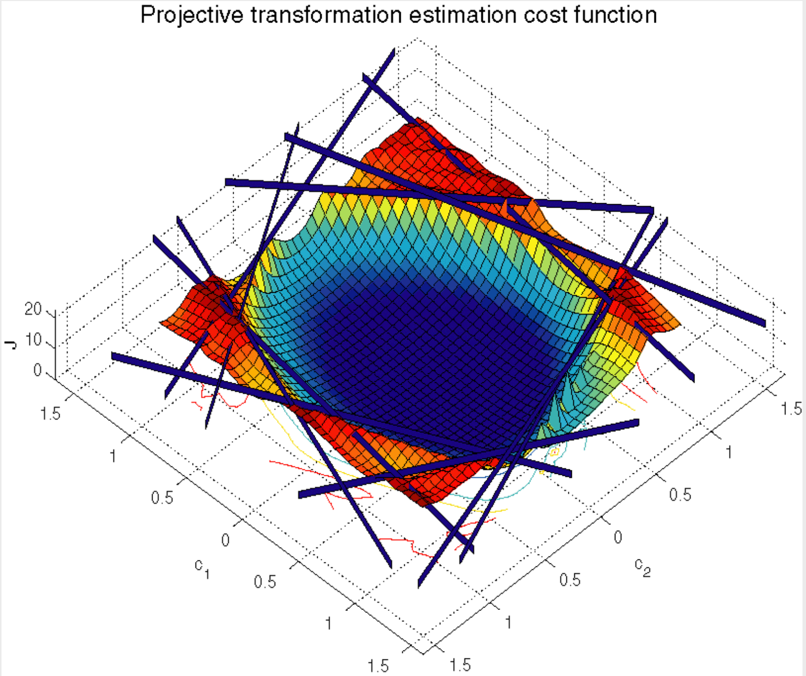
Issues: dimensionality, initialization, complexity

# Reduction to a 2D Problem

- Can re-write normal equations so that optimal $A$, $b$ are functions of optimal $c$.

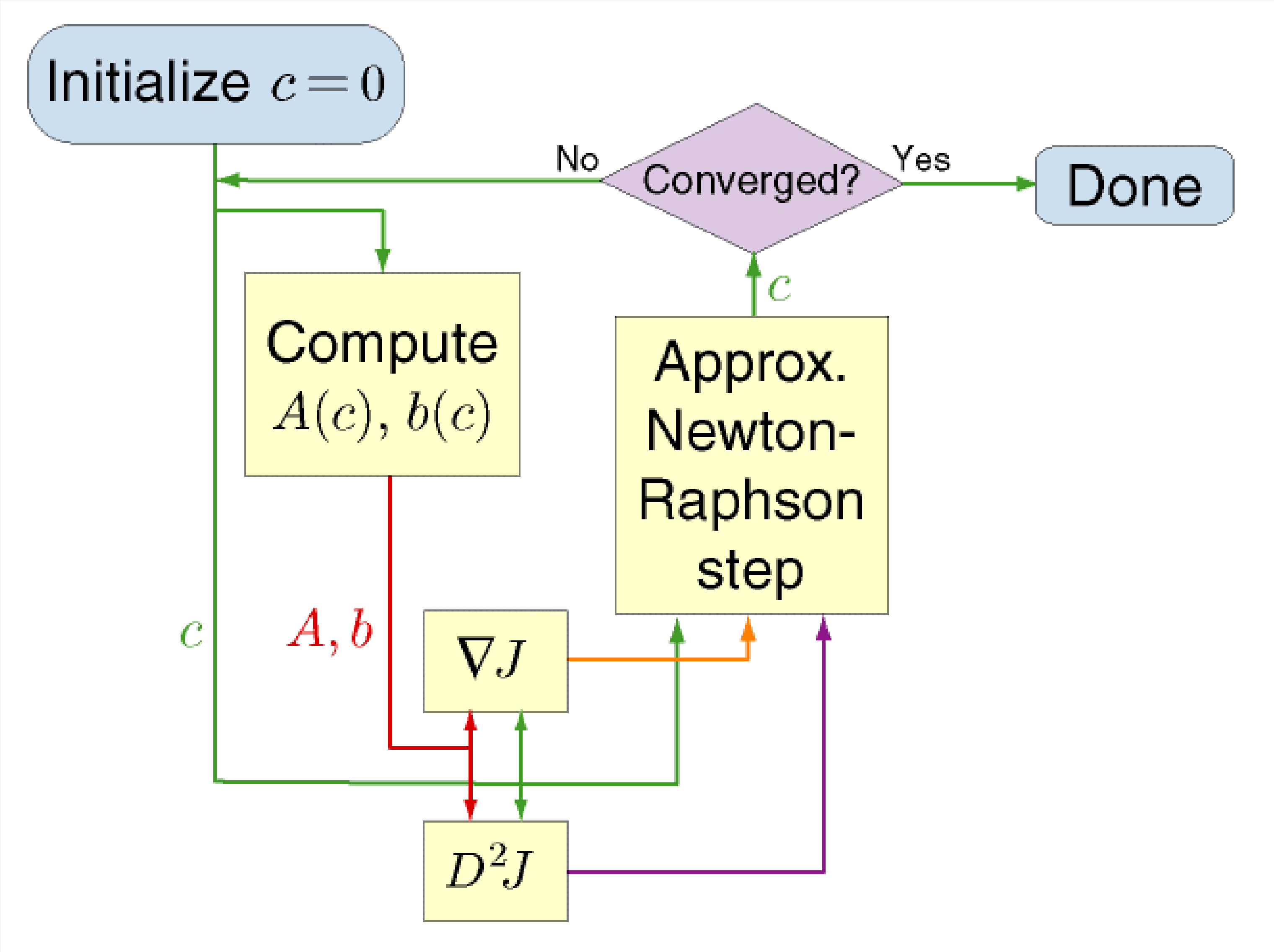
- The least-squares solution lies on a 2-dimensional manifold:

$$M = \left\{ (A, b, c):\ A = A(c),\ b = b(c),\ c \in R^2 \right\}$$



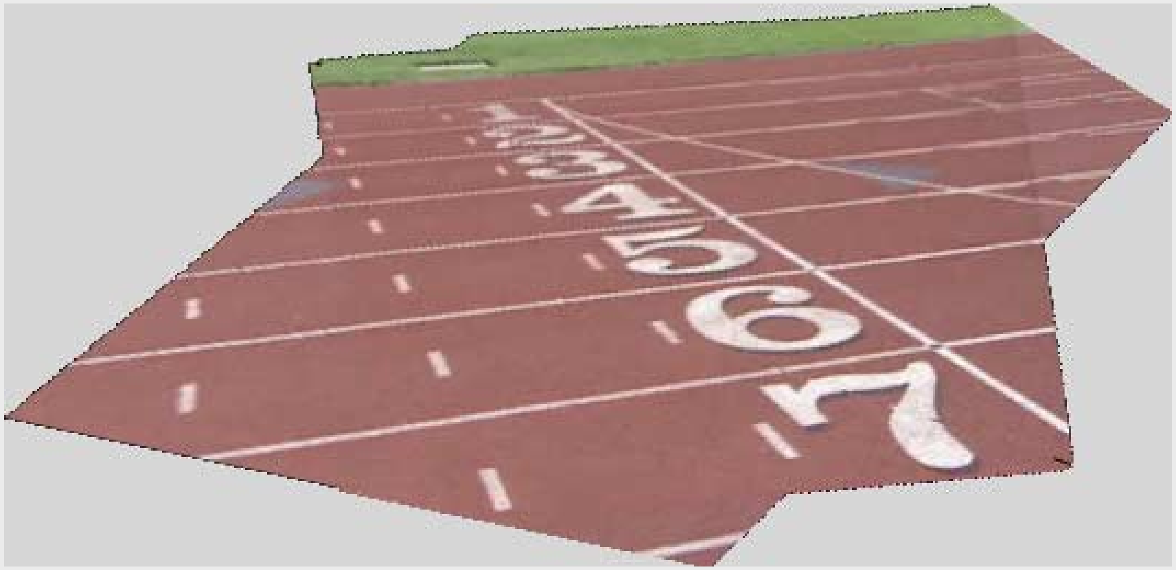

$$\min_c\ J(c) = \sum_{j=1}^{N} \left( w_j' - \frac{A(c)\ w_j + b(c)}{c^T w_j + 1} \right)^T \left( w_j' - \frac{A(c)\ w_j + b(c)}{c^T w_j + 1} \right)$$

# Proposed Algorithm for Minimizing $J$

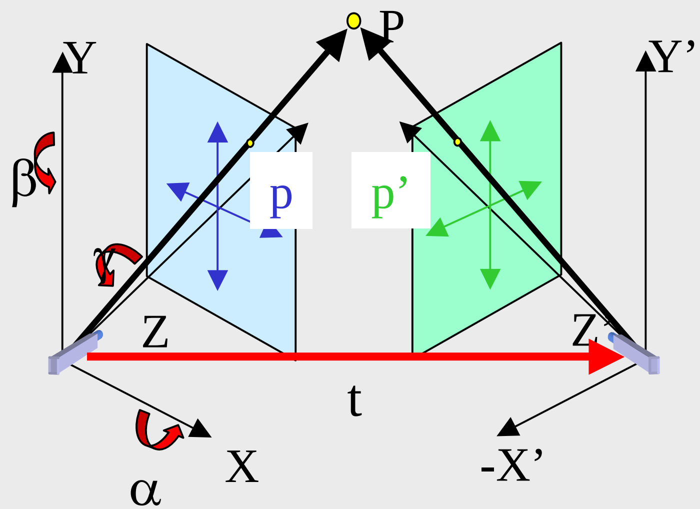


# Image Mosaicking (rotating camera)

# Image Mosaicking (planar scene)



# Obtaining Detailed Signal Correspondence

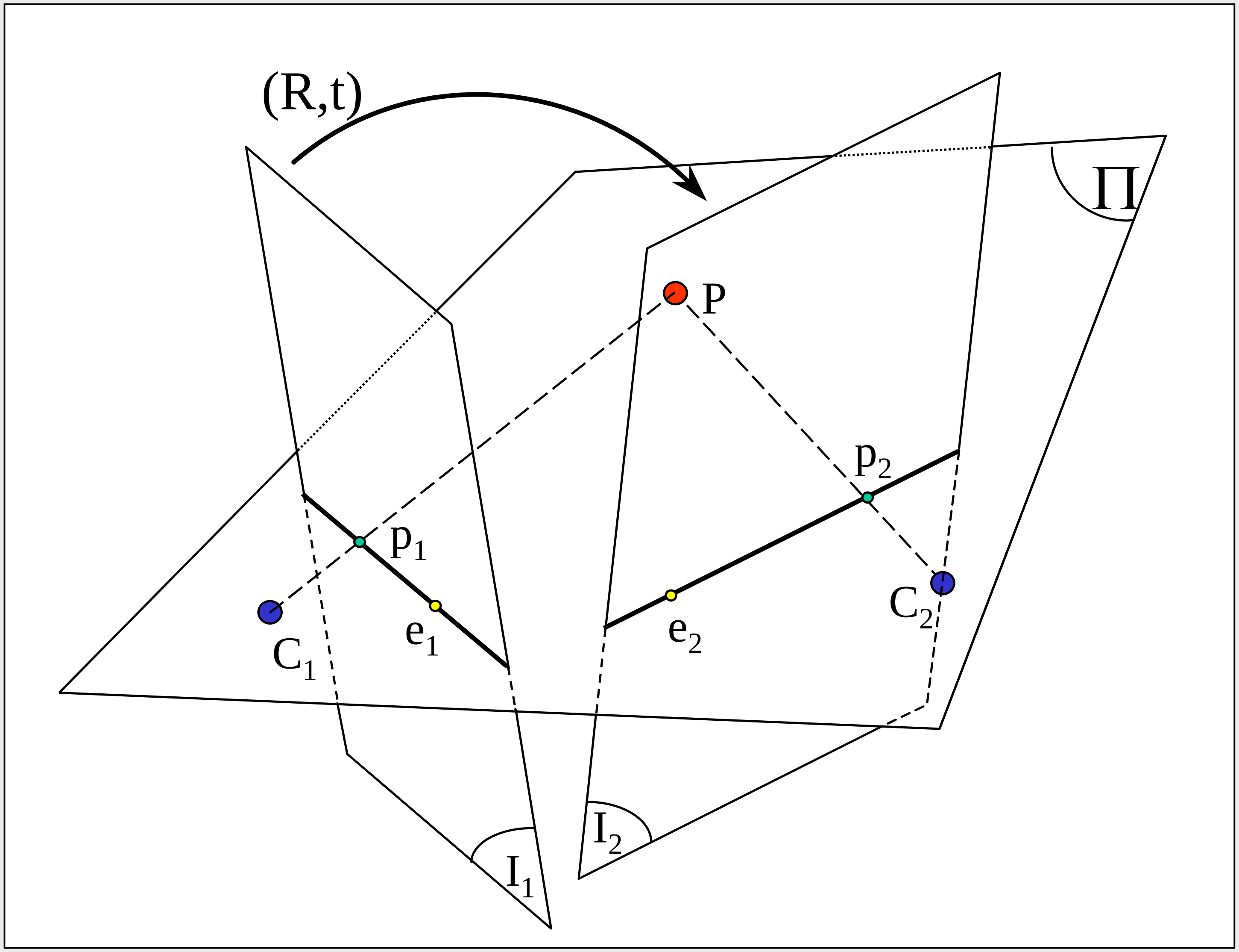



- Detailed signal correspondence depends on both relative sensor geometry and environment (signal sources)
- Classic problem, but more difficult than stereo due to large sensor separation

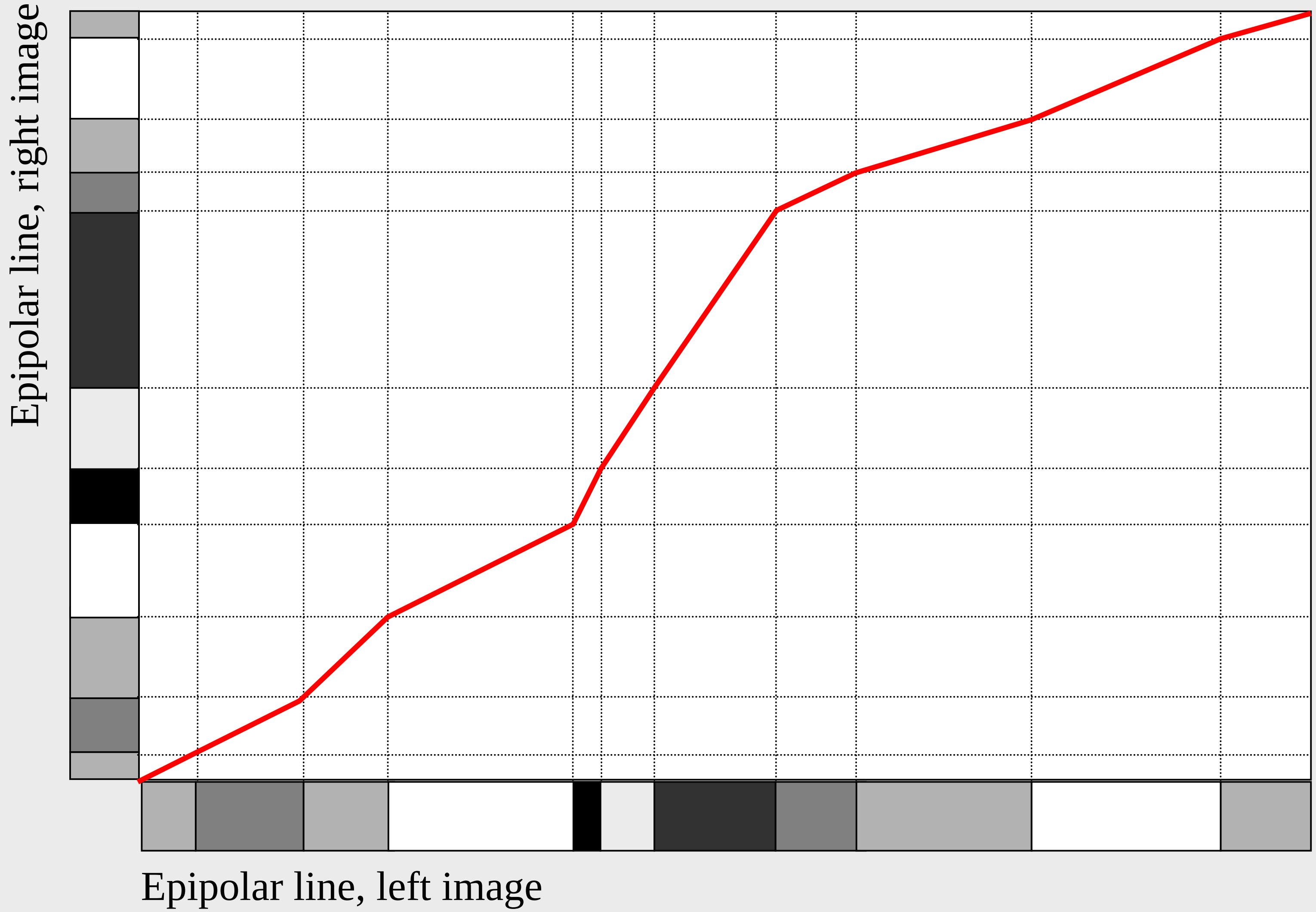
# Estimating Image Correspondence

- Estimate epipolar geometry
- Formulate as finding an optimal path
- Choose interval matching cost function
- Correctly deal with non-monotonicity

# Epipolar Geometry

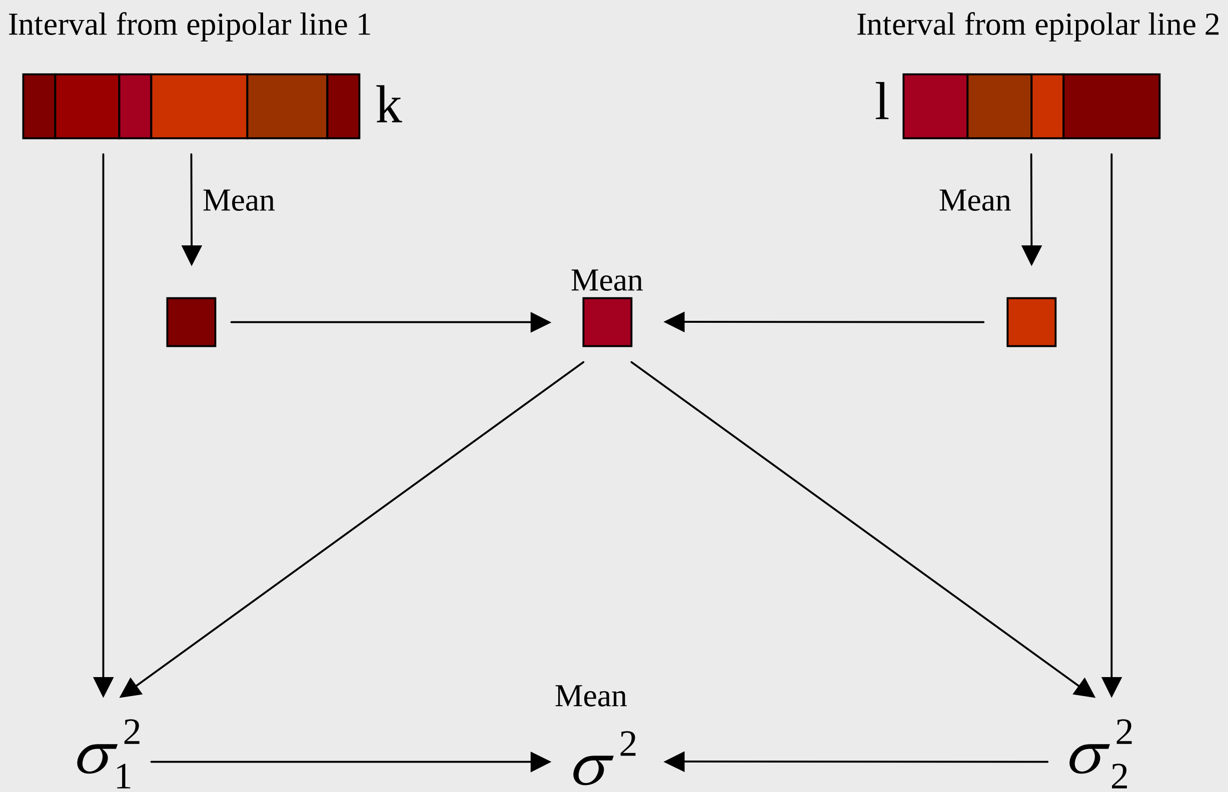


# Matching Graph




Monotonicity allows the use of dynamic programming.

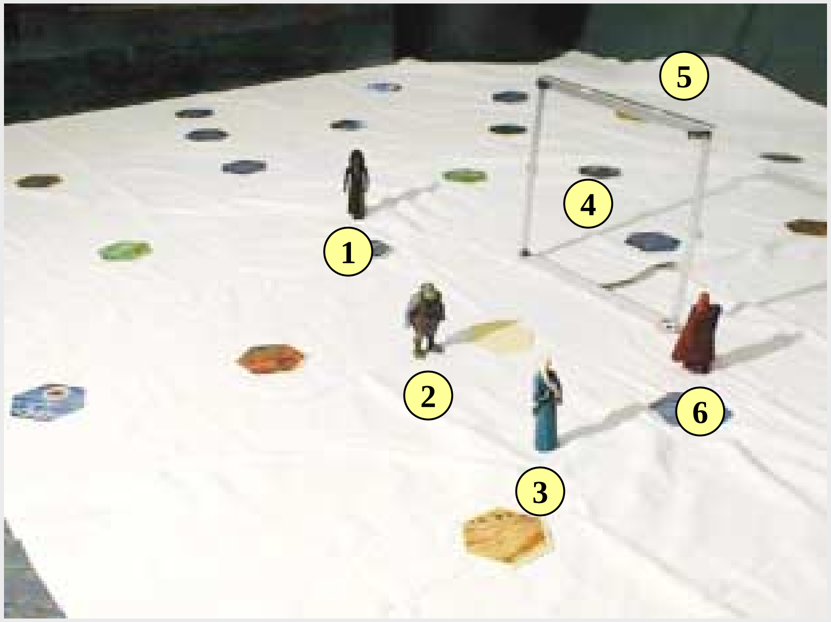
# Interval Matching Cost Function



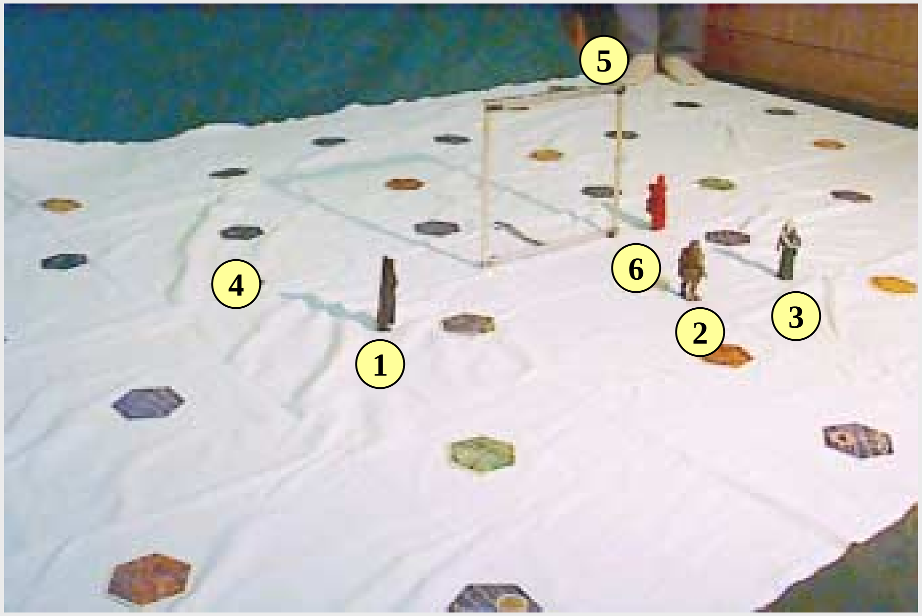

$$C = \sigma^2 \sqrt{k^2 + l^2}$$

Cost is proportional to variance of intensities from the mean, and lengths of the intervals.

# Violations of Monotonicity



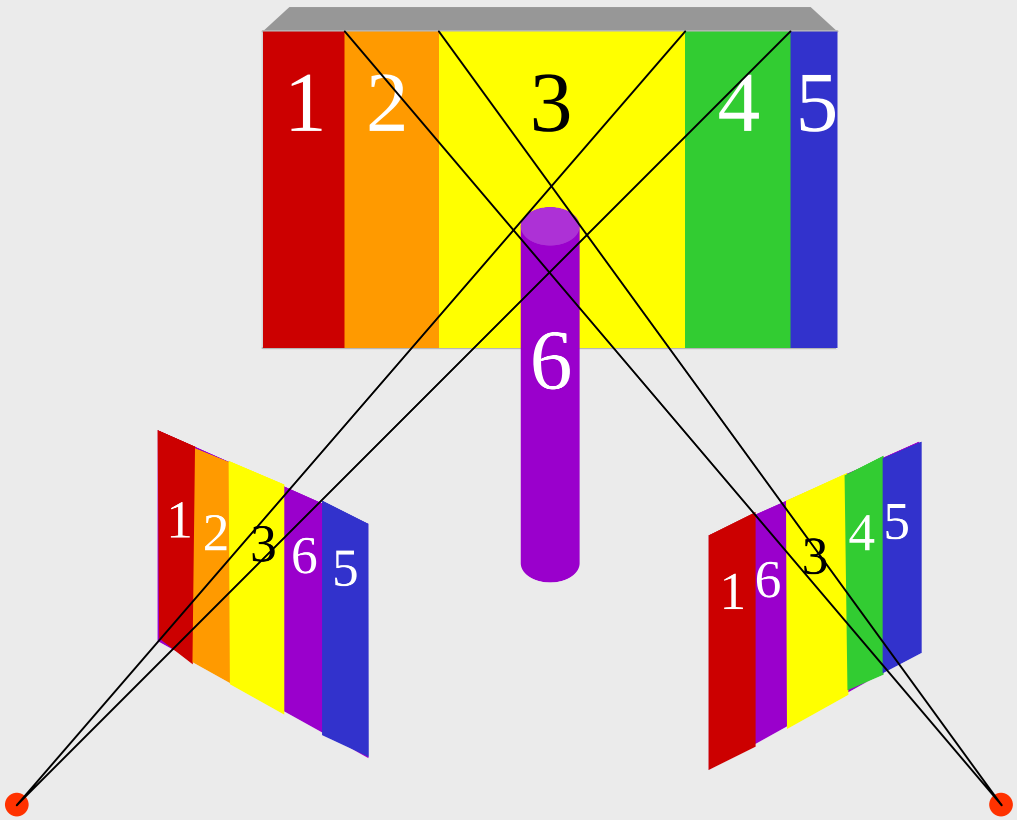

Camera 1

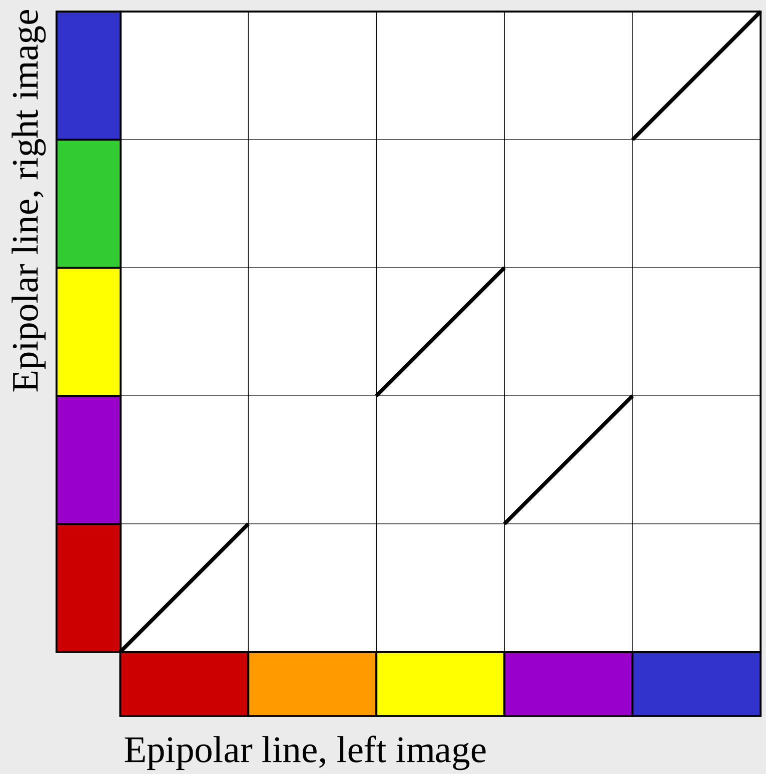



Camera 2

Object arrangement can vary widely between views!

# Occlusions and Monotonicity



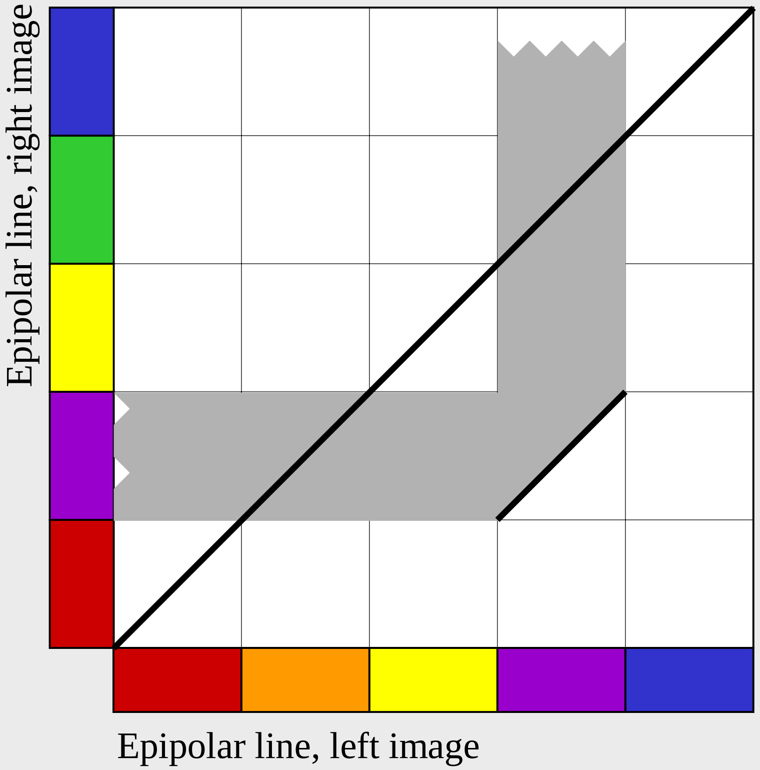

Graph of visible correspondences is:

1) Not monotonic
2) Not continuous

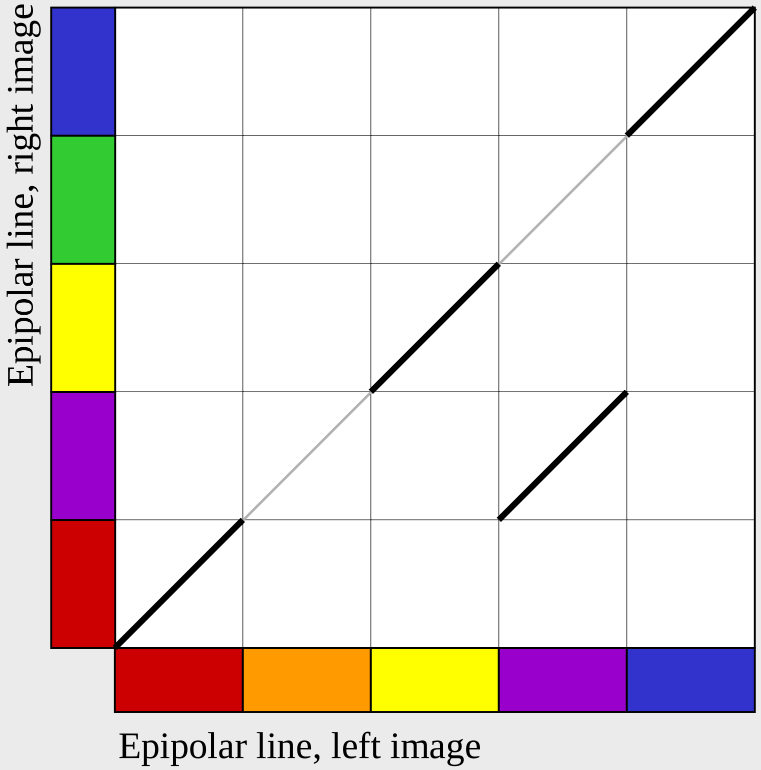
<u>But</u>: is a set of monotonic pieces.

# The Correspondence Graph

The set of all points that are visible in both epipolar lines.

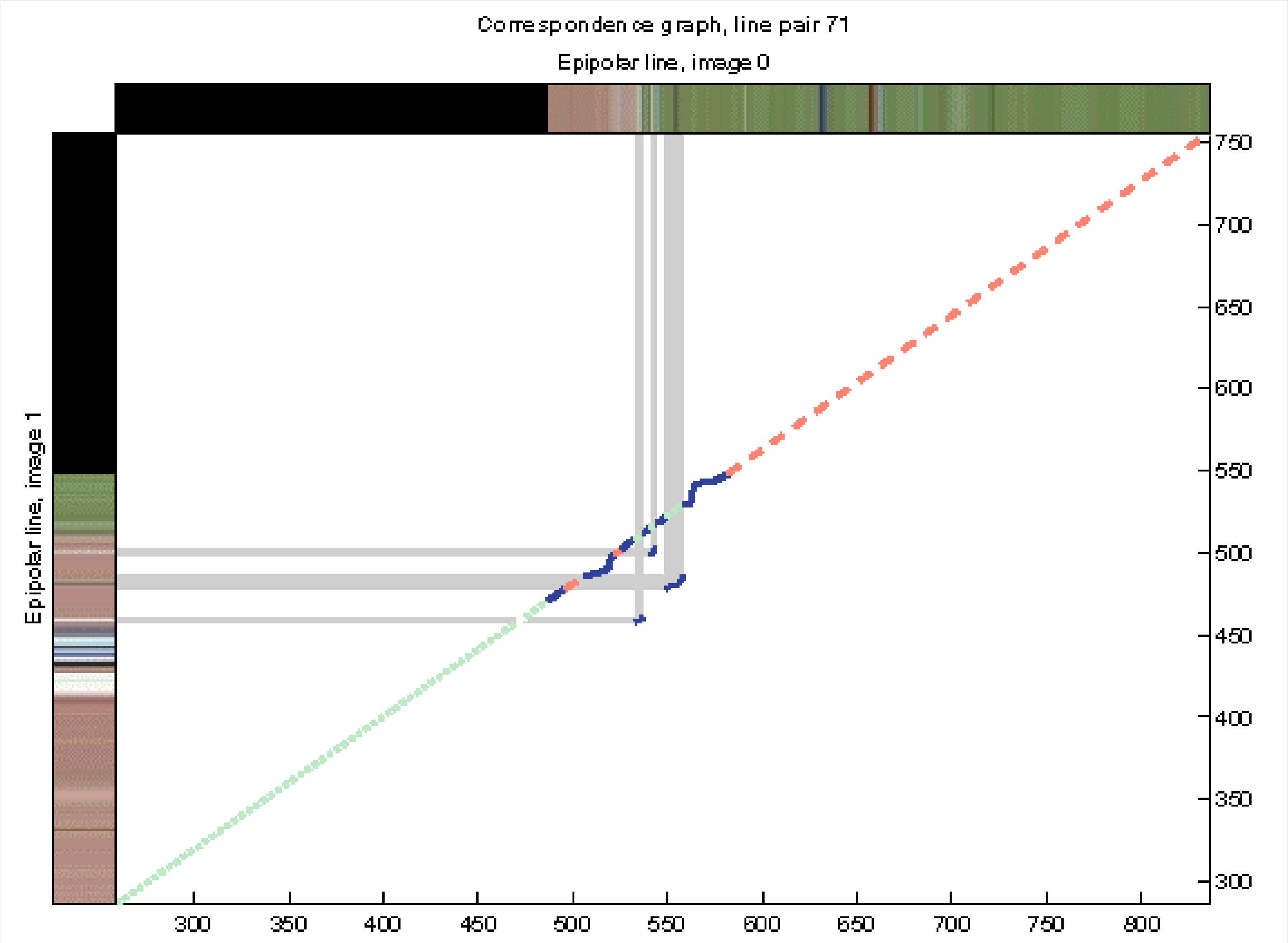



Foreground object +
Background model




Visible correspondences

# A Real Correspondence Graph

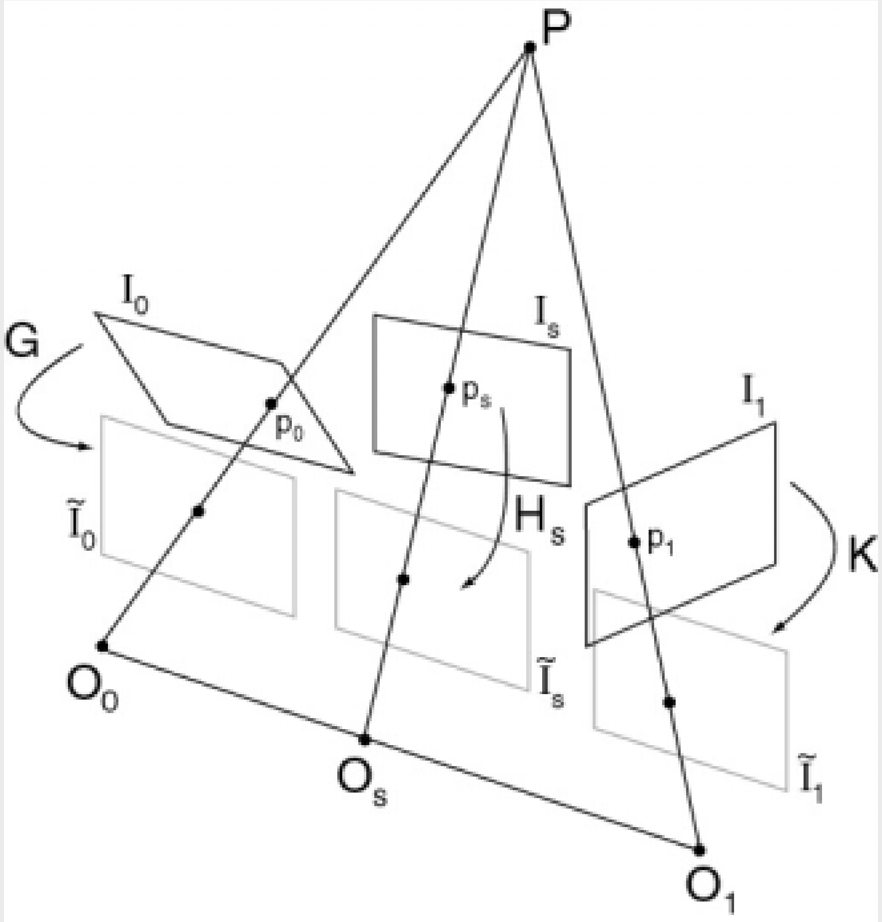



Tells: which regions are visible in both images
which regions are visible in just one image
how to fill in “holes” in correspondence

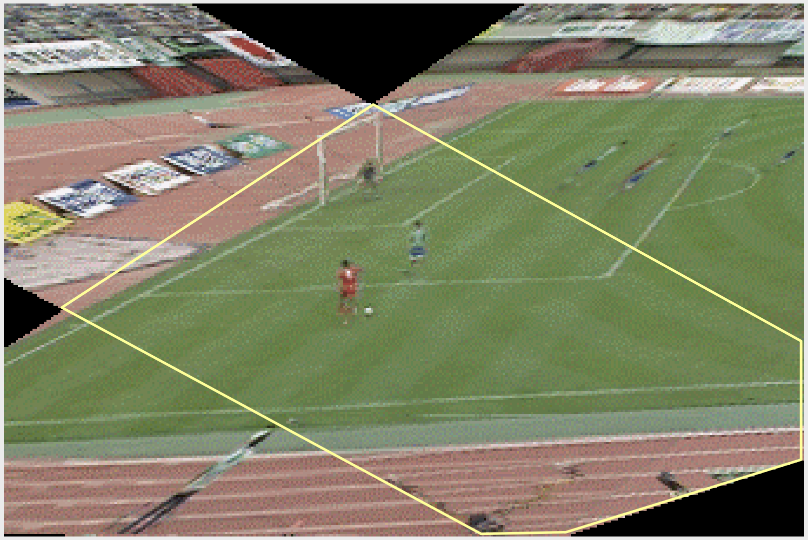
# View Morphing

- Seitz and Dyer, *SIGGRAPH `96*

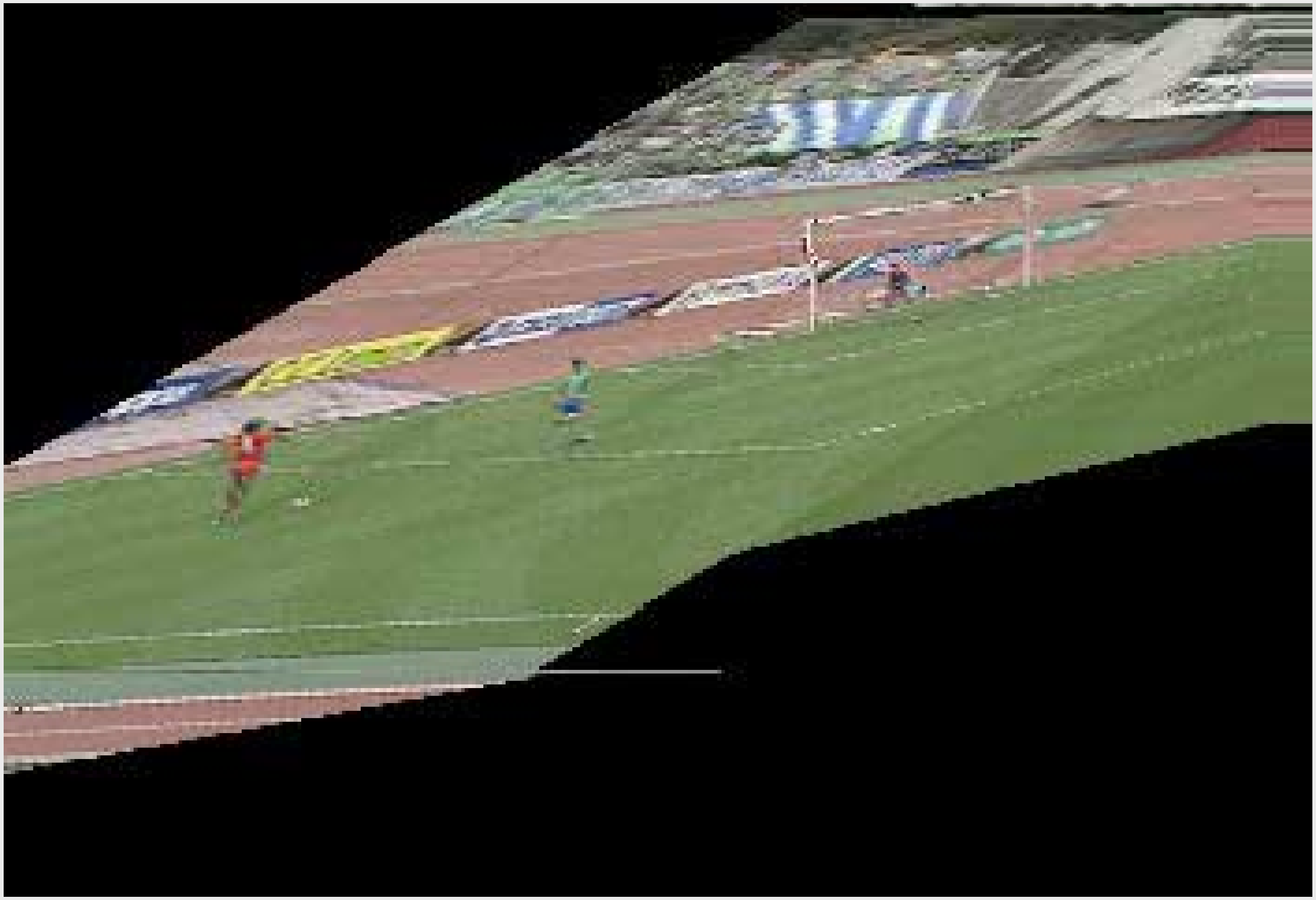
- Rectify image planes
- Virtual camera lies on baseline
- Algorithm depends on pixel-dense correspondence



# A Virtual Image

# Virtual Video from Wide-Baseline Stills



# Recursive Propagation and Fusing Dynamic Streams

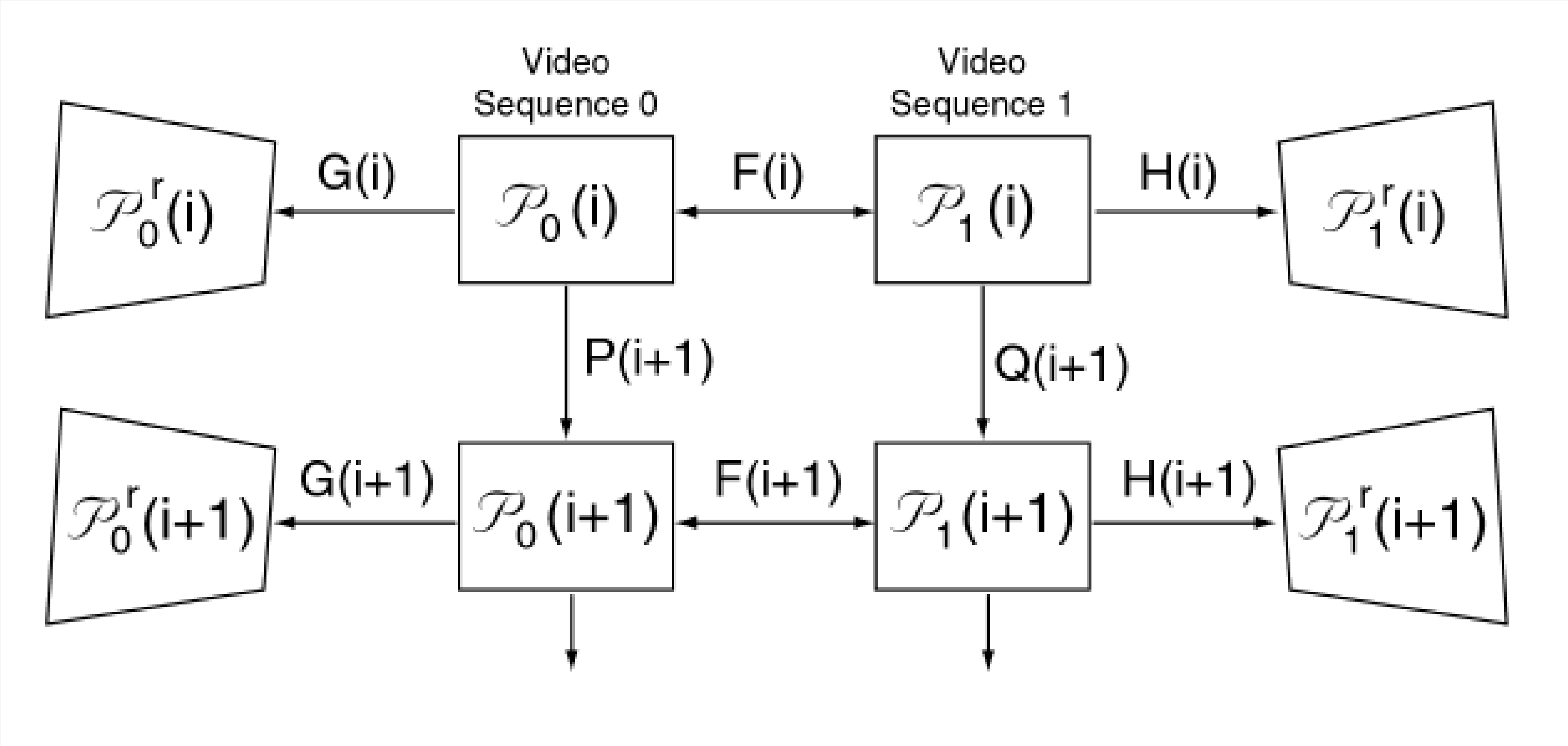
Fusion requires relative geometry and detailed correspondence information.

These are hard, time-consuming problems.

Estimating this information anew at each step is prohibitively expensive.

Approach: A recursive algorithm for the propagation of geometry and correspondence information.

# Relationships Between Images



# Recursive Propagation Equations

$\hat{\chi}(i)$ : an estimate of correspondence between a pair of image planes at time $i$.

$$\hat{\chi}(0|0) = \tilde{\chi}(0)$$

$$\hat{\chi}(i+1|i) = T^{i+1}(\hat{\chi}(i|i))$$

$$\hat{\chi}(i+1|i+1) = M^{i+1}(\hat{\chi}(i+1|i))$$

# Time Update

Let $(p_0, p_1)$ be a correspondence in $\mathrm{P}_0(i)$ x $\mathrm{P}_1(i)$.

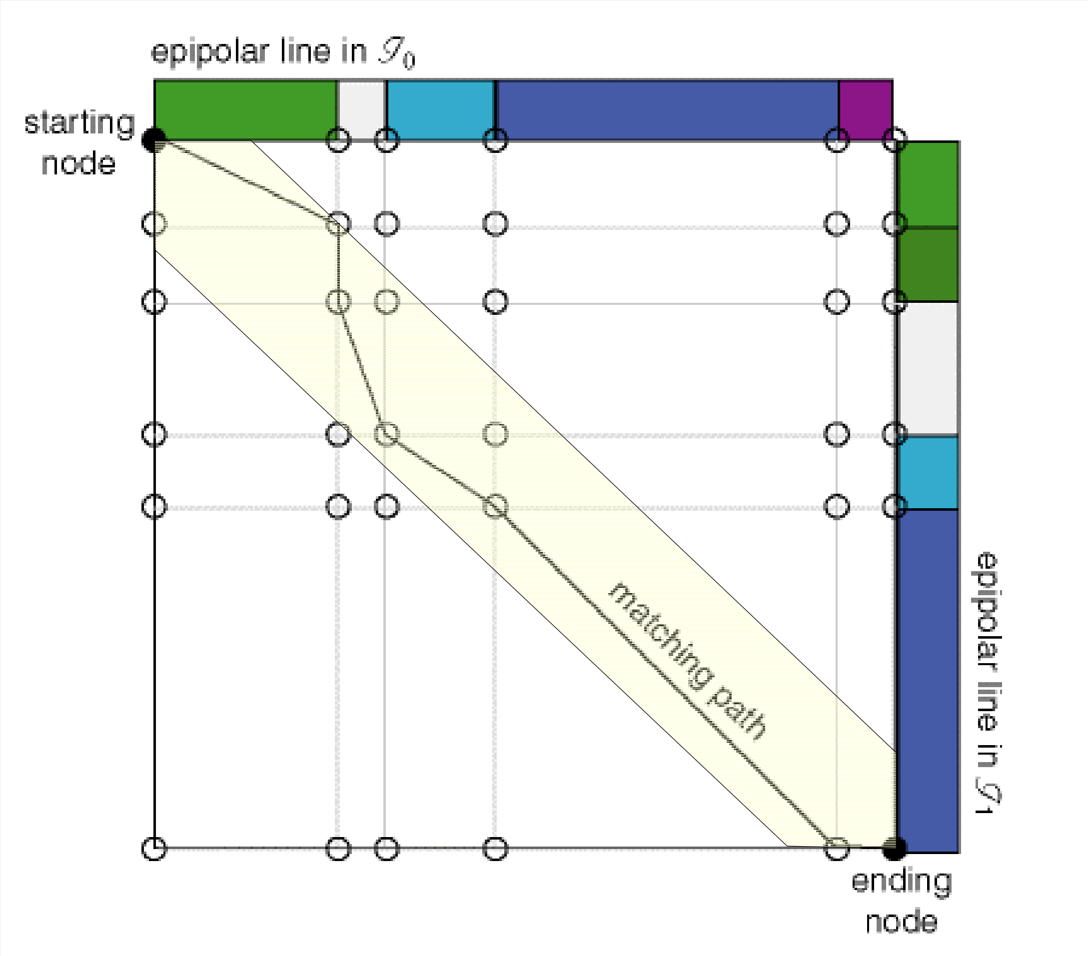
The time update is:

$$T^{i+1}(p_{0,} p_1) = (P(i+1)p_0 \;,\; Q(i+1)p_1)$$

In practice, we use an approximation $\hat{T}^{i+1}$.

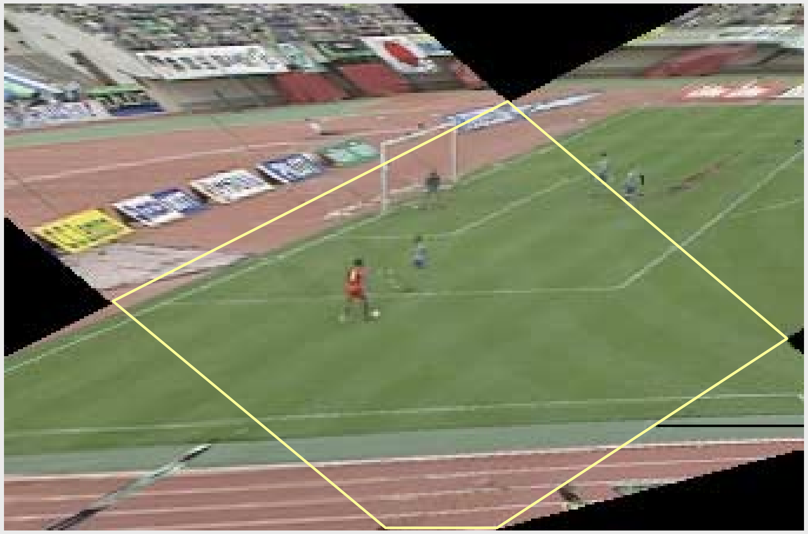
Using appropriate rectifying projective transformations, the time update becomes the identity.

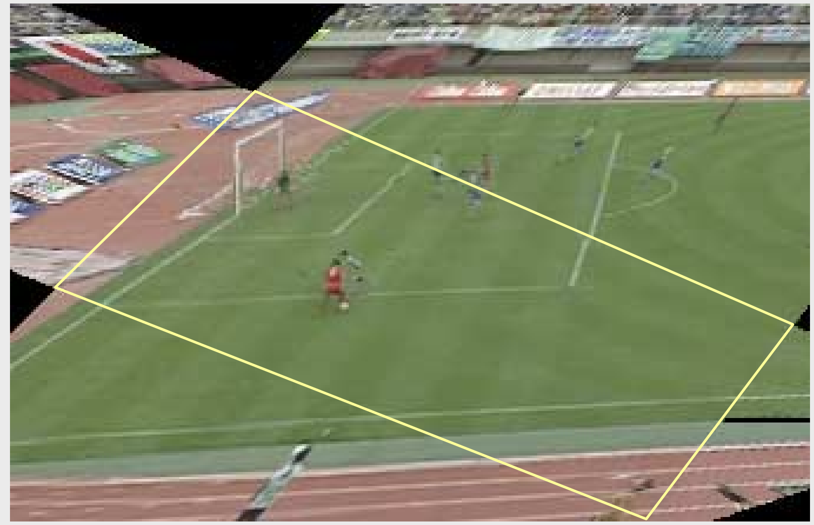
# Measurement Update



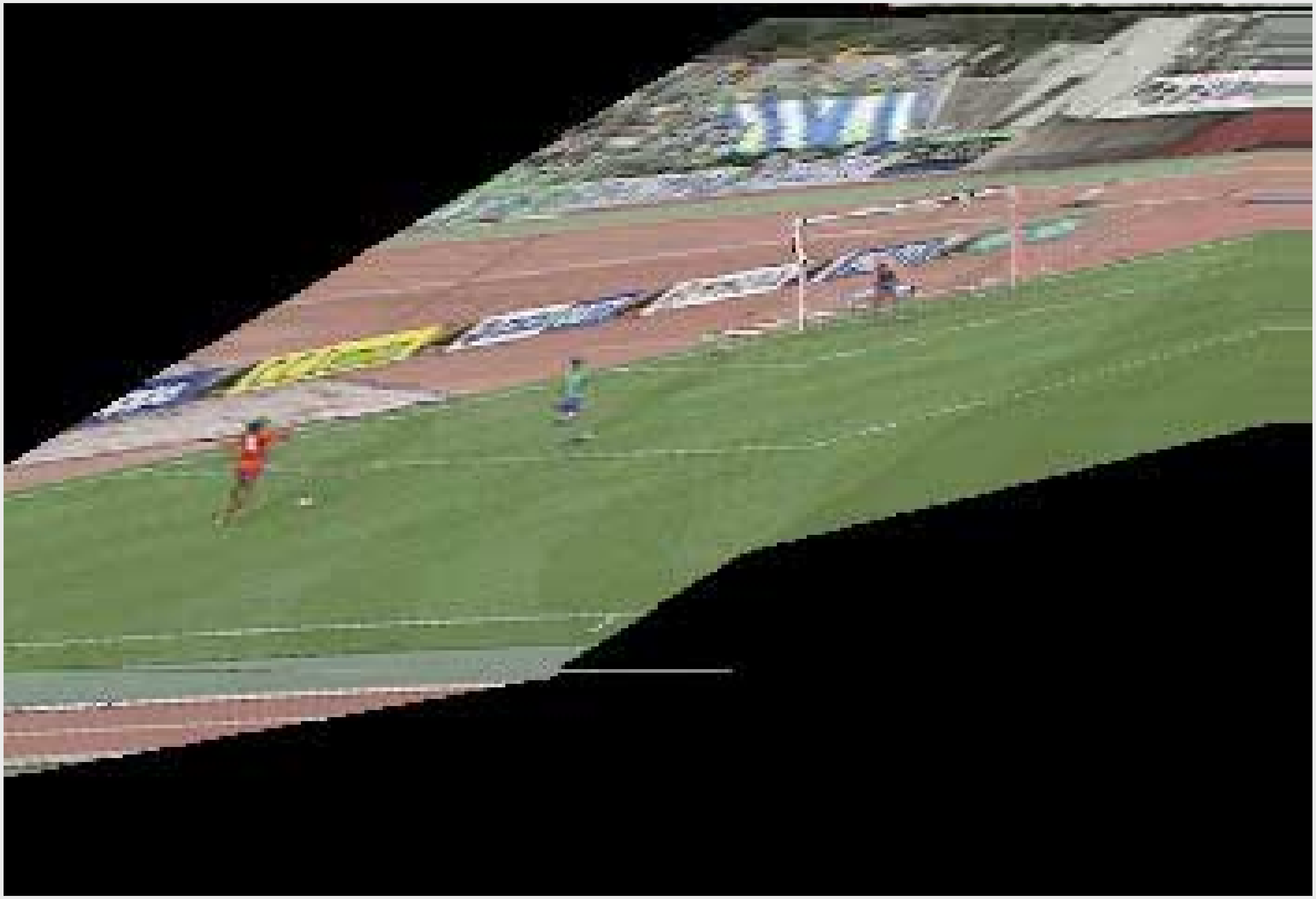

Dynamic programming confined to a neighborhood of the time-updated estimate.

# Virtual Video






# Virtual Video of Dynamic Scene from Two Video Streams



# Summary and Future Work

- Methodology for fusing uncalibrated myopic sensors to obtain global/joint information
- Estimating relative sensor geometry
- Detailed sensor correspondence
- Recursive propagation and fusing dynamic sensor streams

- Limitations with less resolution, etc.
- Deal with other sensor types
- Joint consideration of many sensors
- Effects of limited computation, communication

[redacted]