# Computational Photography

**Prof. Feng Liu**

**Spring 2022**

http://www.cs.pdx.edu/~fliu/courses/cs510/

**05/17/2022**

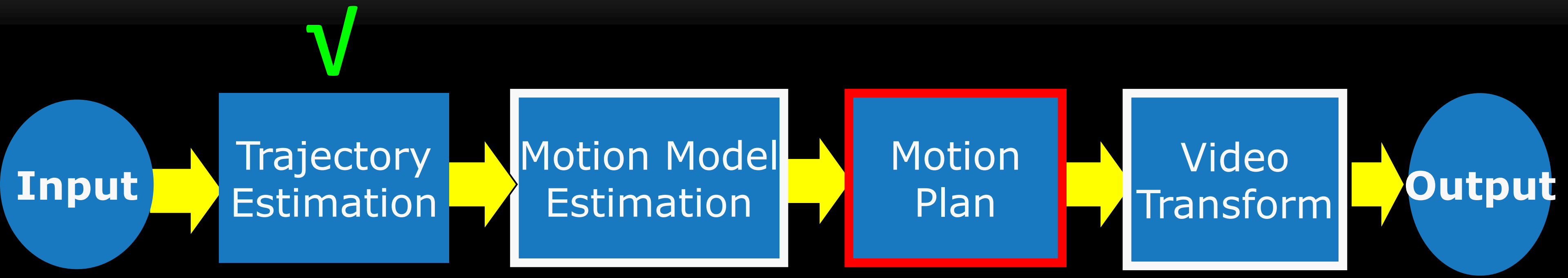
# Last Time

- Video Stabilization
  - Video stabilization pipeline

# Today

- Video Stabilization
  - 3D Video Stabilization
  - Subspace Video Stabilization

# Traditional 2D Video Stabilization



# Motion Plan

$$\mathbf{S}_t = \sum_{i \in N_t} \mathbf{T}_t^i * G, \text{where } \mathbf{T}_t^i = \prod_{j=i}^{t} \mathbf{T}_j$$

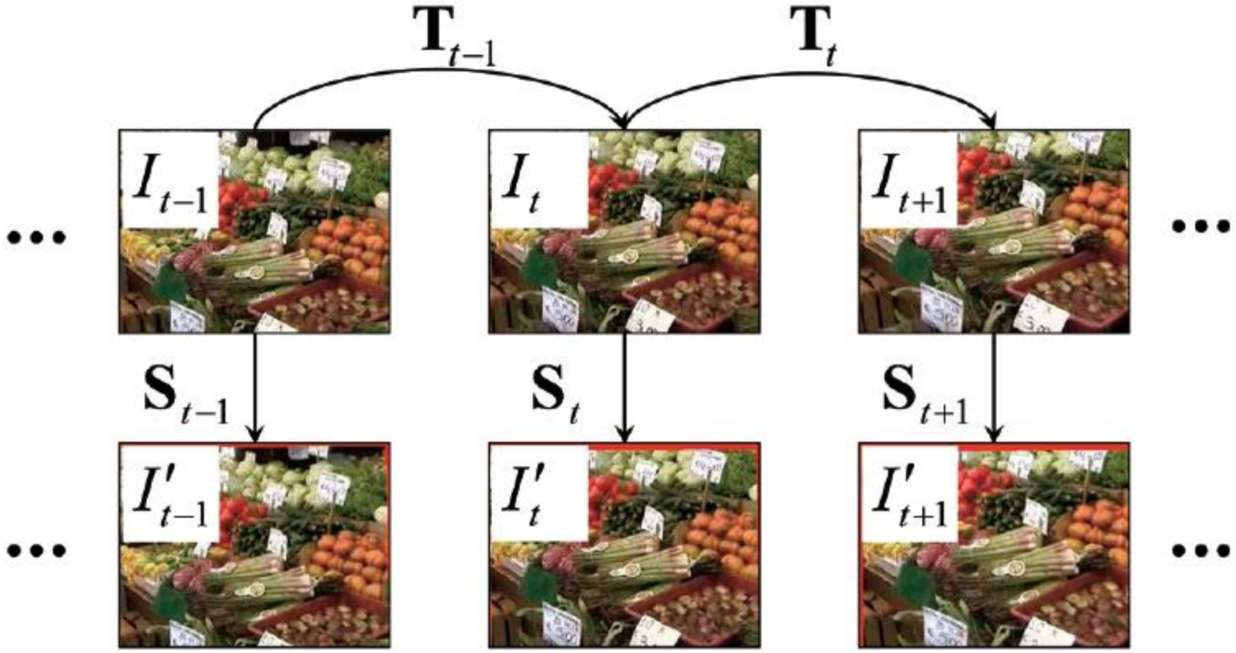





Image courtesy: Matsushita et al. 06

# Traditional 2D Video Stabilization Result

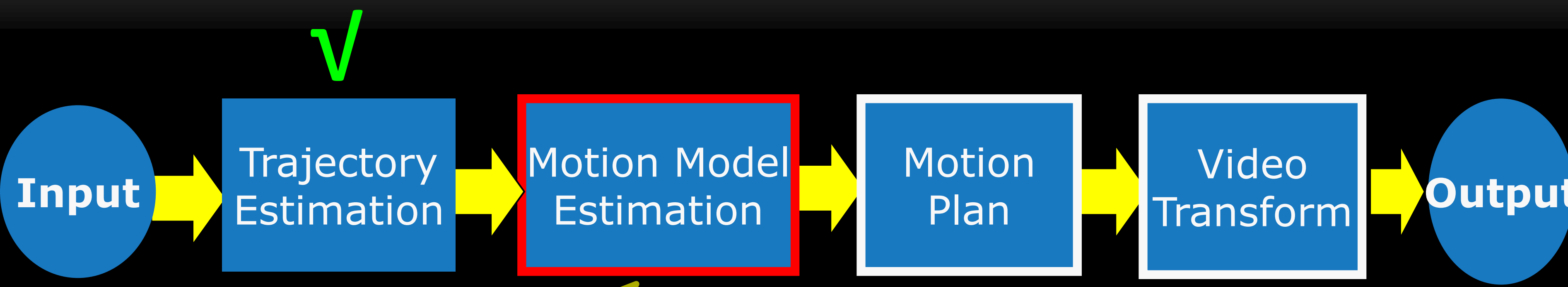
# Limitations

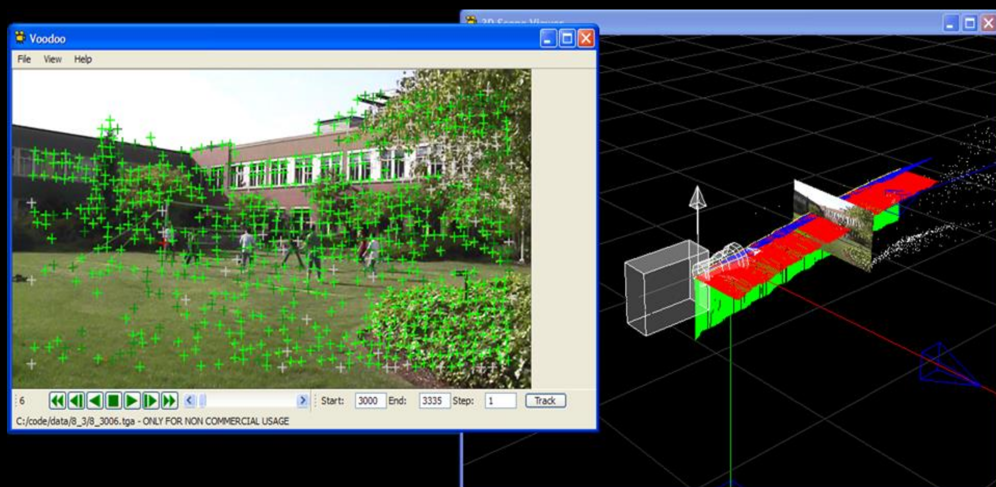
- No knowledge of actual 3D camera path, so cannot control desired motion directly
- Homography cannot model 3D camera motion and scene structure

# 3D Video Stabilization

- Non-metric image-based rendering for video stabilization [Buehler et al. 01]
- Image-based rendering using image-based priors [Fitzgibbon et al. 05]
- Using photographs to enhance videos of a static scene [Bhat et al. 07]

# 3D Video Stabilization

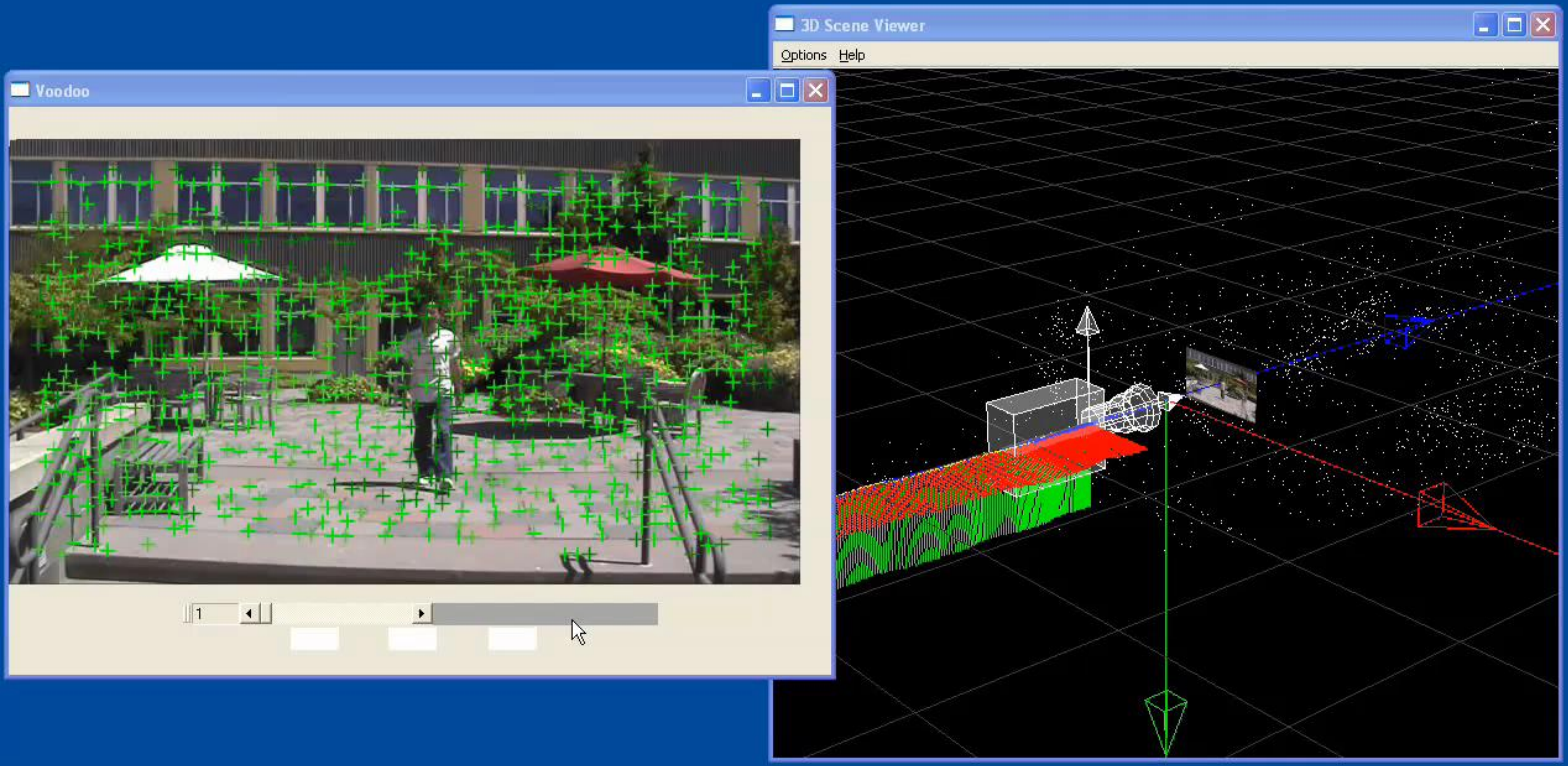



3D reconstruction via *structure from motion*

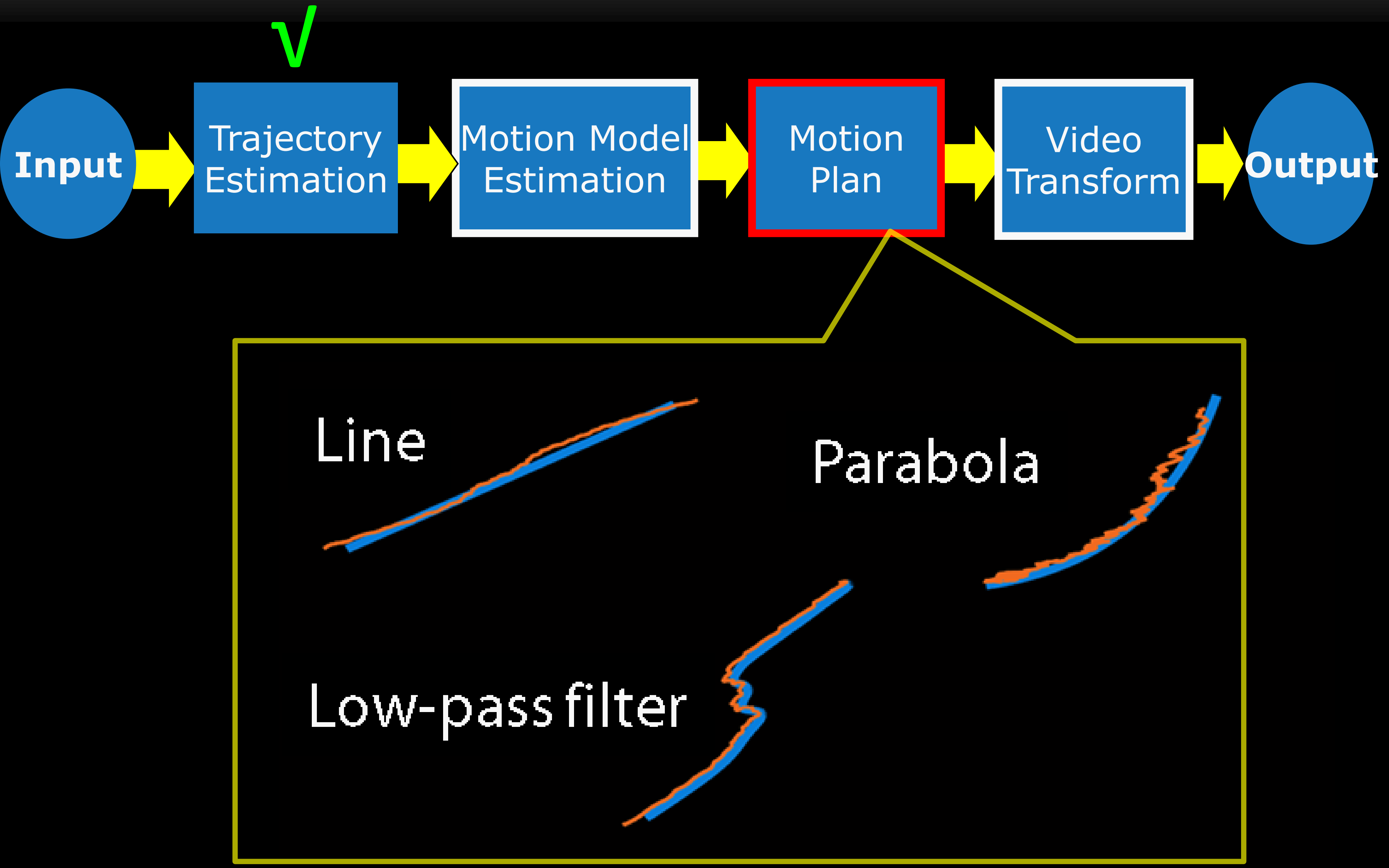
Voodoo Camera Tracker `(http://www.digilab.uni-hannover.de)`

# Structure from Motion

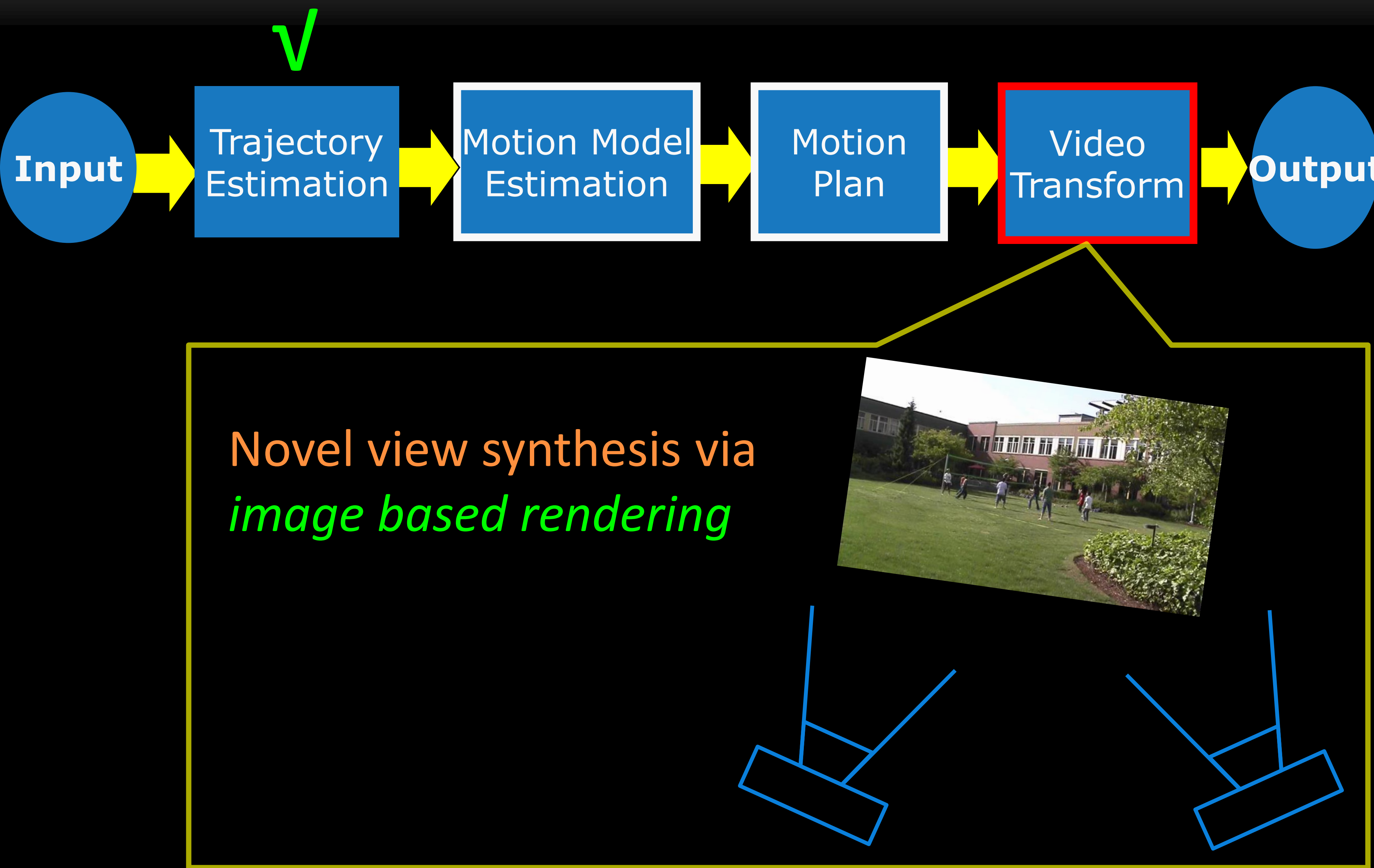



Voodoo Camera Tracker (http://www.digilab.uni-hannover.de)

# 3D Video Stabilization



# 3D Video Stabilization



# Novel View Synthesis by Image based Rendering

Unstructured lumigraph rendering [Buehler et al. 01]

# Content-preserving warps based 3D video stabilization

F Liu, M Gleicher, H Jin, A Agarwala. Content-preserving warps for 3D video stabilization, SIGGRAPH 2009

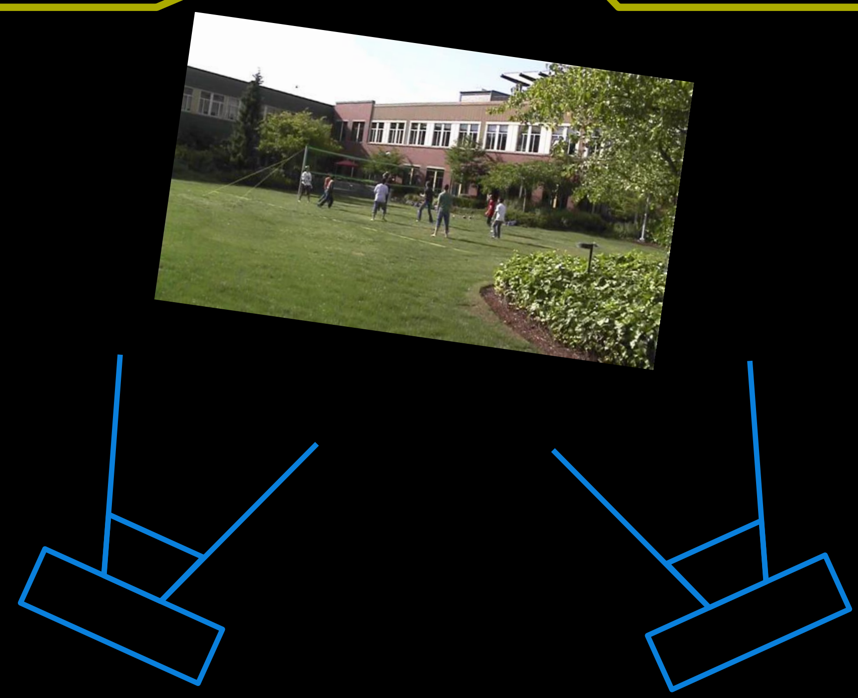
# 3D Video Stabilization

√

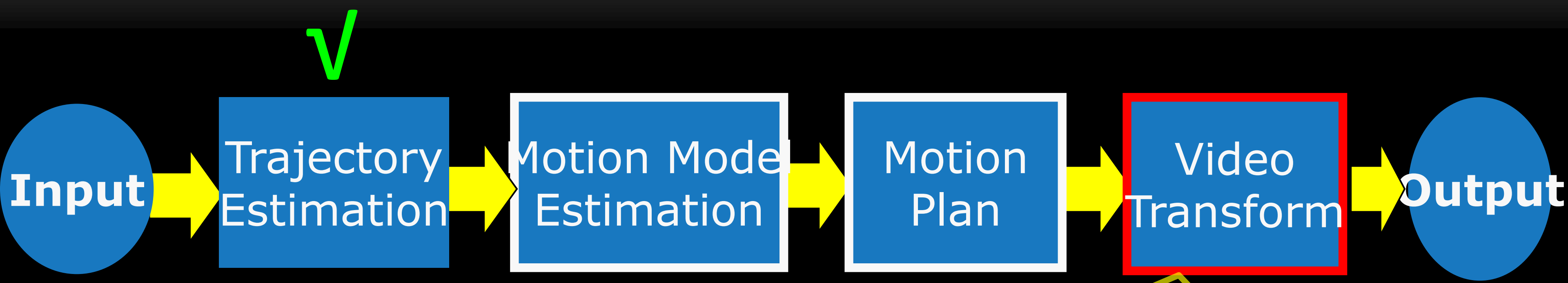
Input → Trajectory Estimation → Motion Model Estimation → Motion Plan → Video Transform → Output

Novel view synthesis

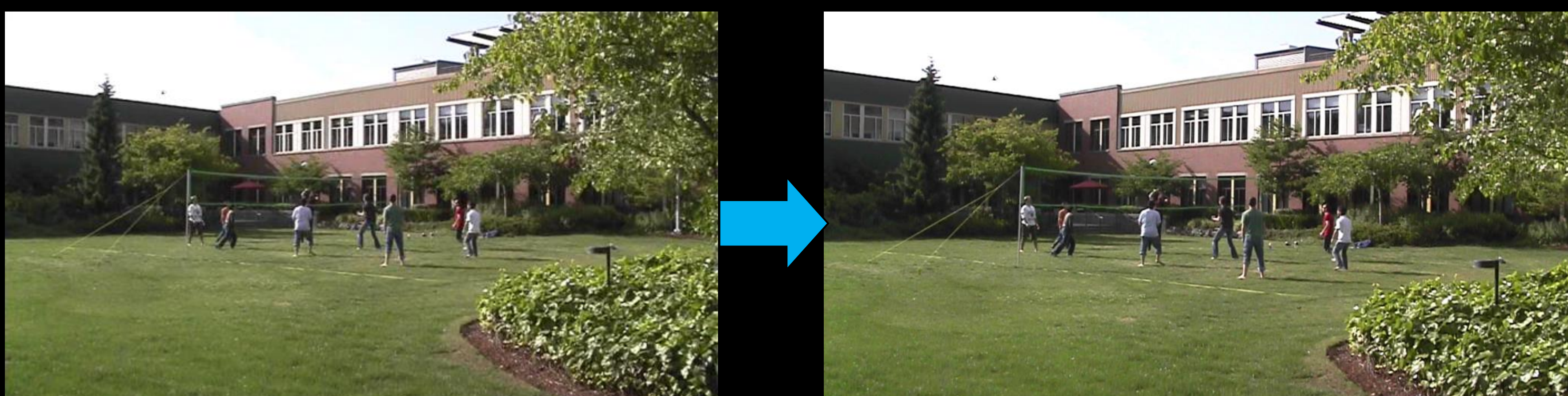
image based rendering



# Temporal Constraint




Our method for novel view synthesis



One input frame

One output frame

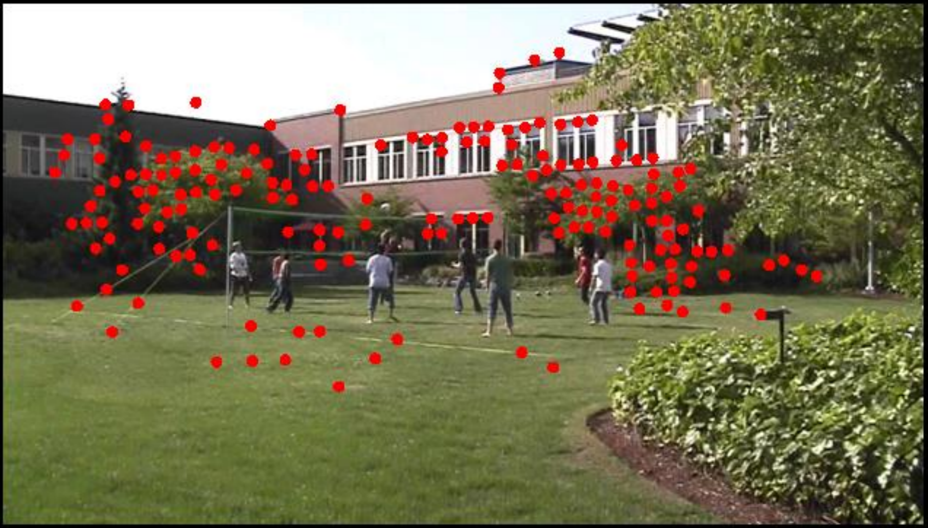
# Novel View from One Frame

- A Series of Vision Challenges!
  - Segment out layers
  - Determine depth
  - Shift and re-composite layers
  - Fill holes
- Cannot achieve accurate dis-occlusions, non-Lambertian reflection, etc.

# Human Perception

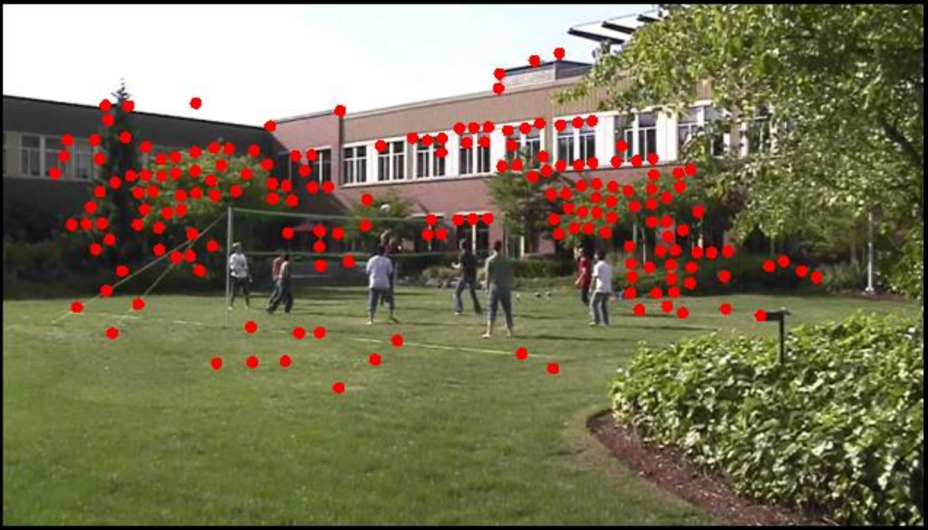
- Viewpoint shifts will be small
- Aim for perceptual plausibility rather than accurate novel view synthesis
  - Move salient content along stabilized paths
  - No noticeable artifacts

# Problem Setup

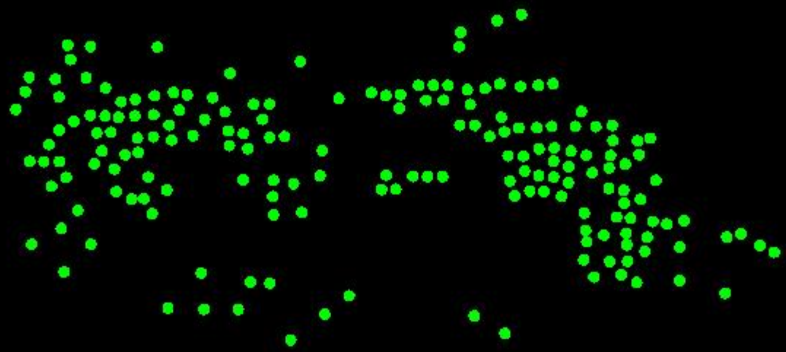


input frame and points

# Problem Setup

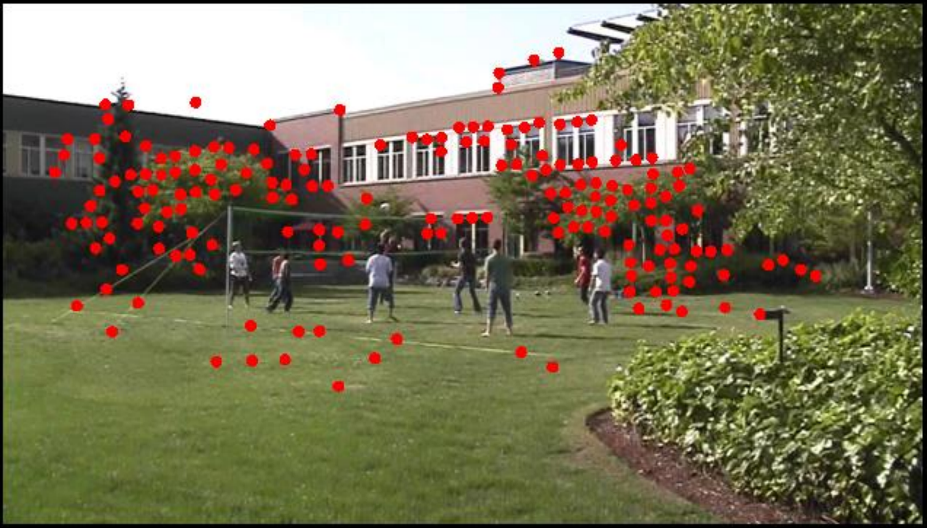


input frame and points



output points

# Problem Setup



input frame and points

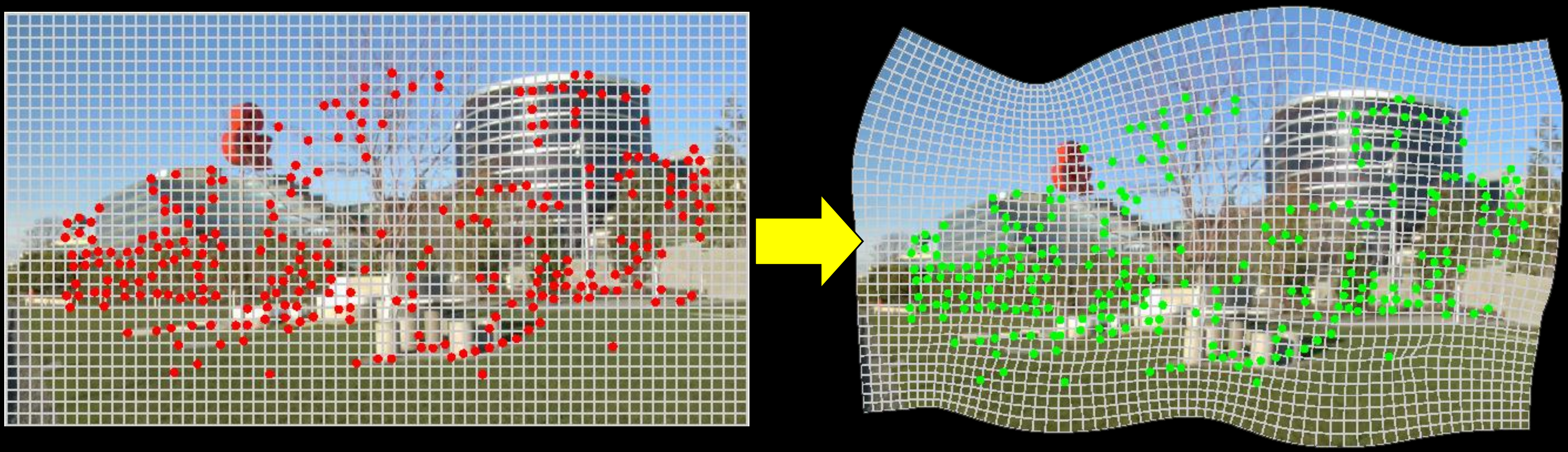
output frame

# Option 1: Scattered Data Interpolation

# Option 2: Full-frame Warping with Homography

# A Less Successful Result

# Our Approach: Content-preserving Warping

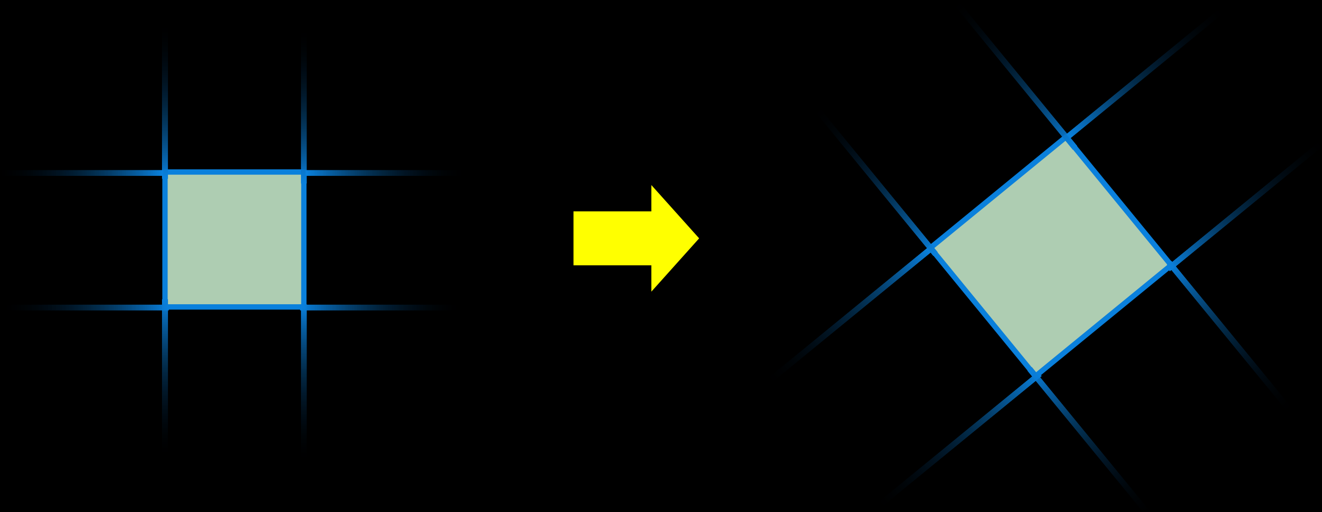


Warp each input frame to create the output frame by least-squares minimization

- ✓ Data term: Soft, sparse displacement constraint
- ✓ Smoothness term: Local similarity transformation constraint

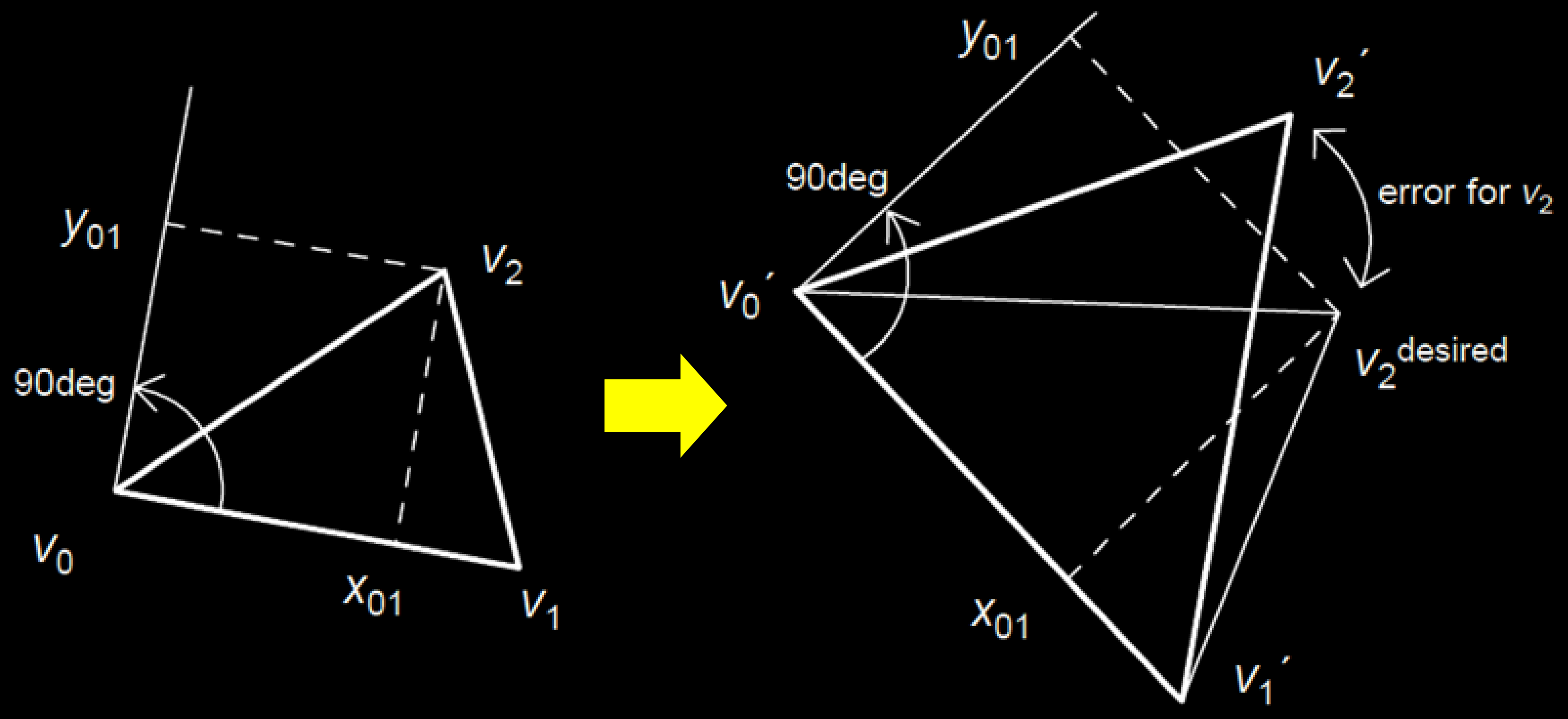
# Smoothness Term: Minimize Visual Distortion

## Local similarity transformation constraint



# Smoothness Term: Minimize Visual Distortion

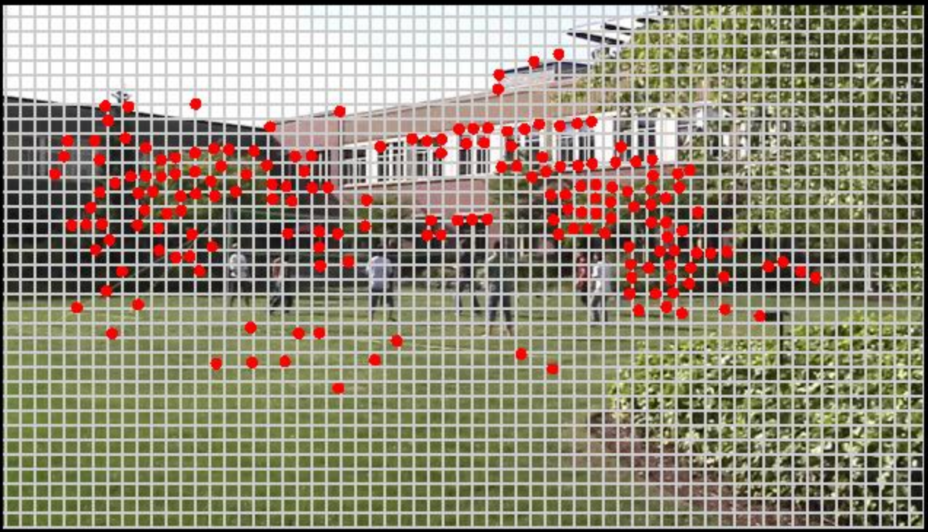
## Local similarity transformation constraint




[Igarashi et al. 05]

# Saliency Weight

Concentrate distortion to non-salient regions



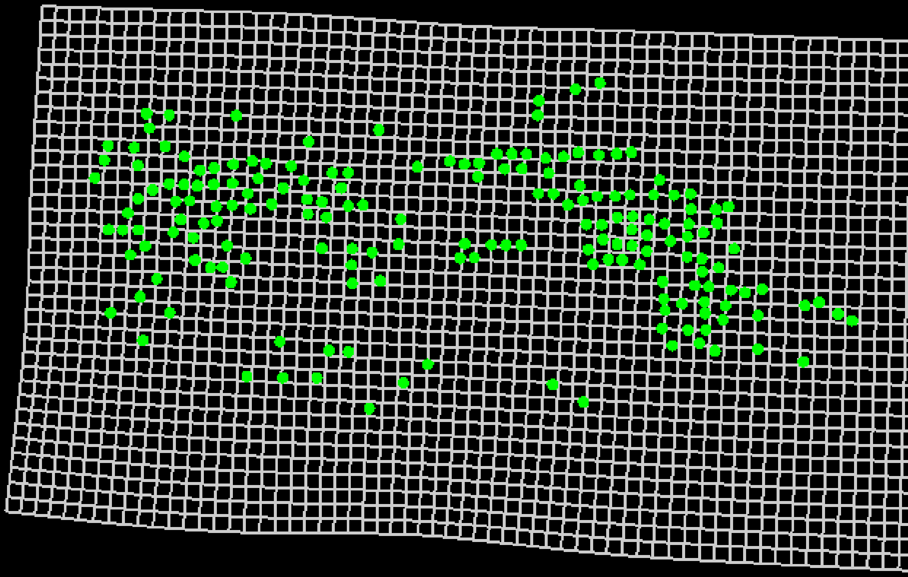
Input

Visual saliency map
[Itti et al. 99]

**Visual saliency**: “the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention” from [Itti 07]

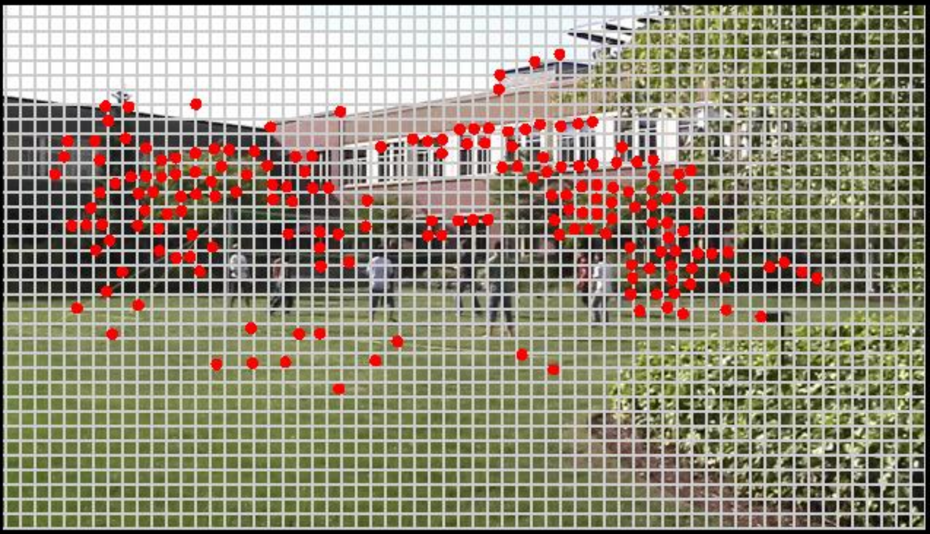
# Content-Preserving Warping

Input



Output

# Content-Preserving Warping



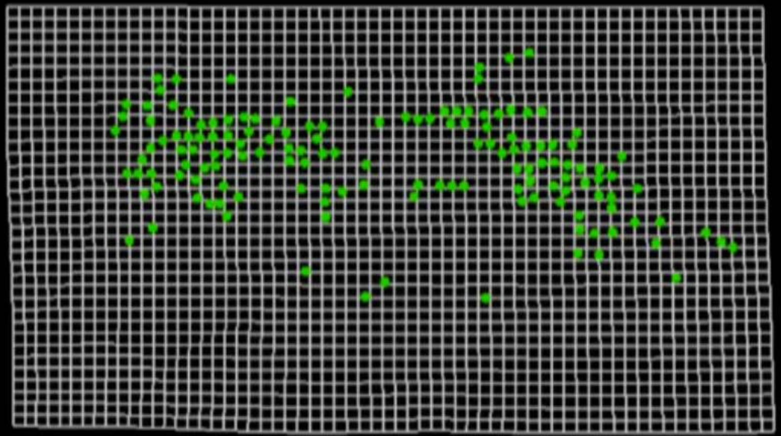
Input

Output

texture mapping [Shirley et al. 2005]

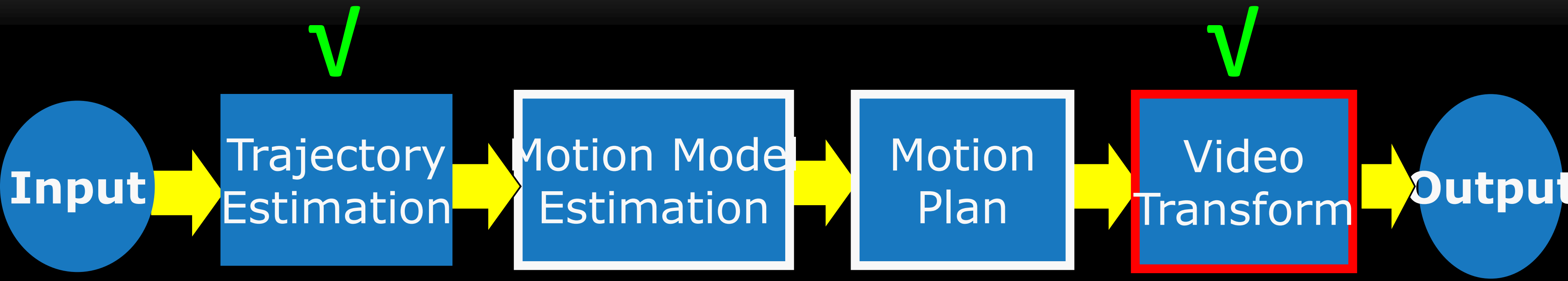
# Content-Preserving Warping

Grid mesh
& points



Output

# Video Stabilization Pipeline




## Further Improvement

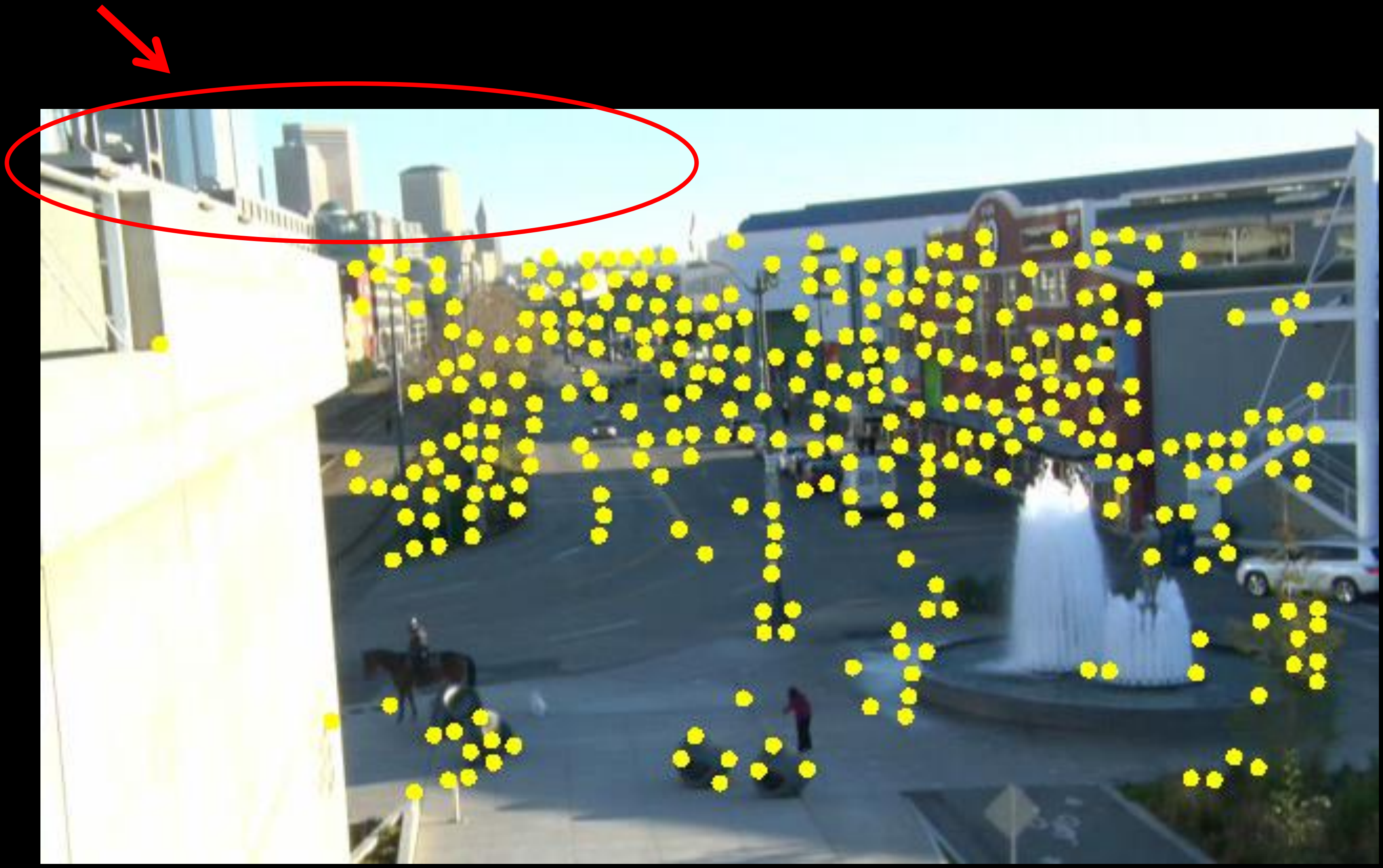
I. Scene points are sometimes poorly distributed

II. The set of feature point changes over time

# Poor distribution of scene points

# Poor distribution of scene points

# No feature points



# Pre-Warping


Input frame

# Pre-Warping


Pre-warping

# Pre-Warping


Pre-warping + content-preserving warping

# Pre-Warping

**Method:**

Pre-warp input using a best-fitting homography

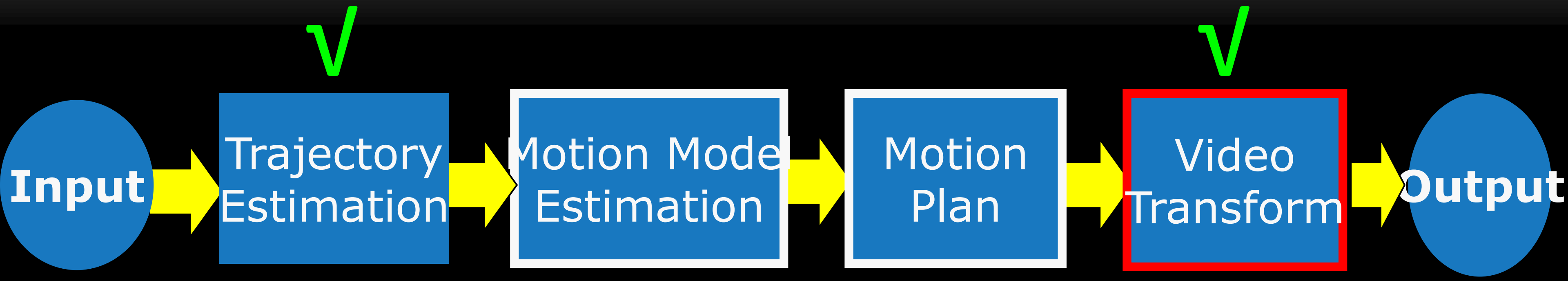
**Result:**

√ Regions with sufficient feature points:
Content-preserving warping dominates

√ Regions without sufficient feature points:
Pre-warping gives a good approximation

Result of content-preserving warping with pre-warping

# Video Stabilization Pipeline

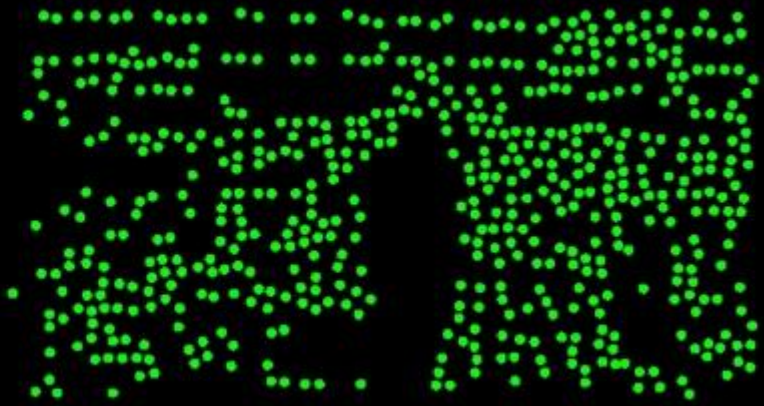



## Further Improvement

I. Scene points are sometimes poorly distributed

II. The set of feature point changes over time

# Temporal Coherence

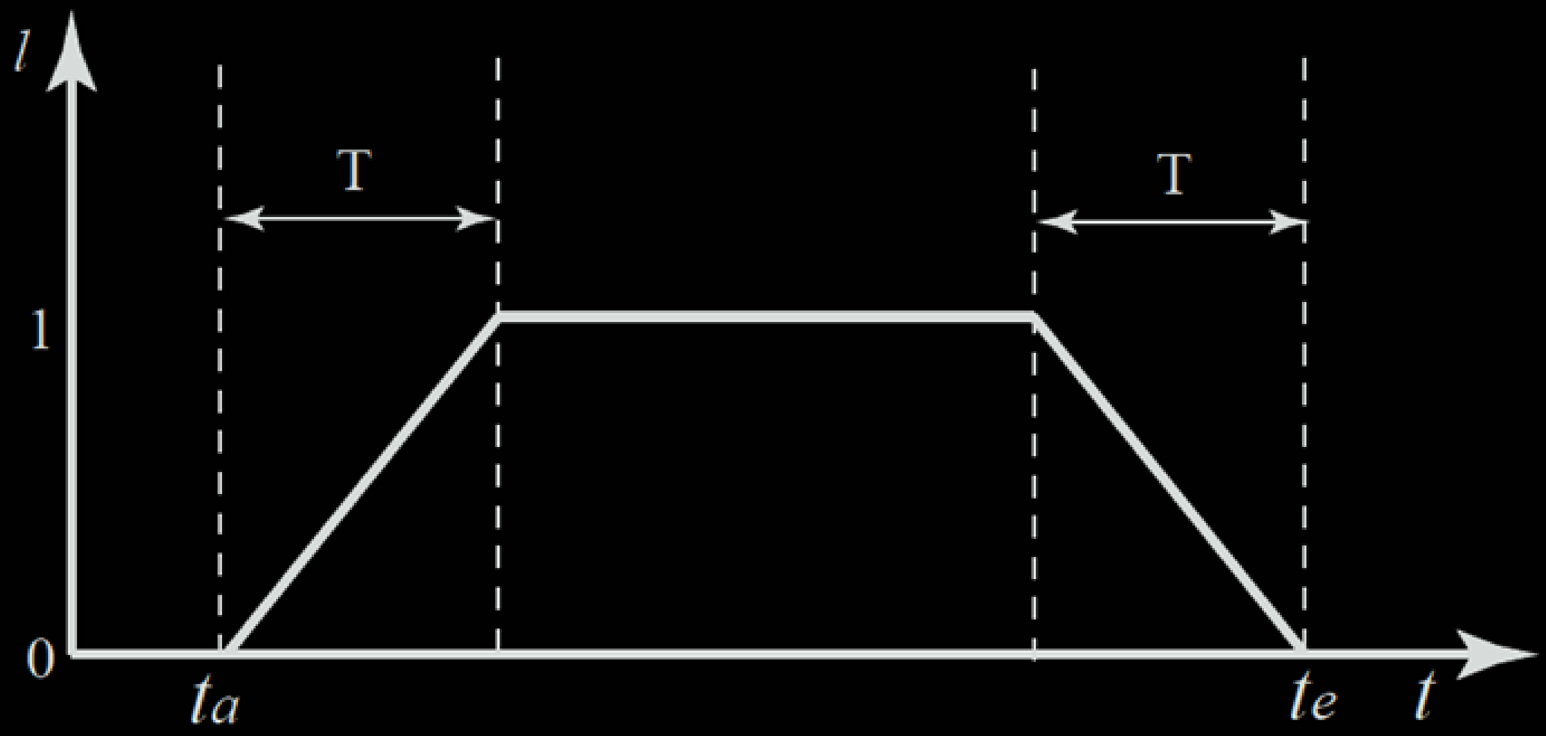


Input video & points

Output points

# Temporal Coherence

# Temporal Coherence




Fade-in/out the weight of the data constraint over time

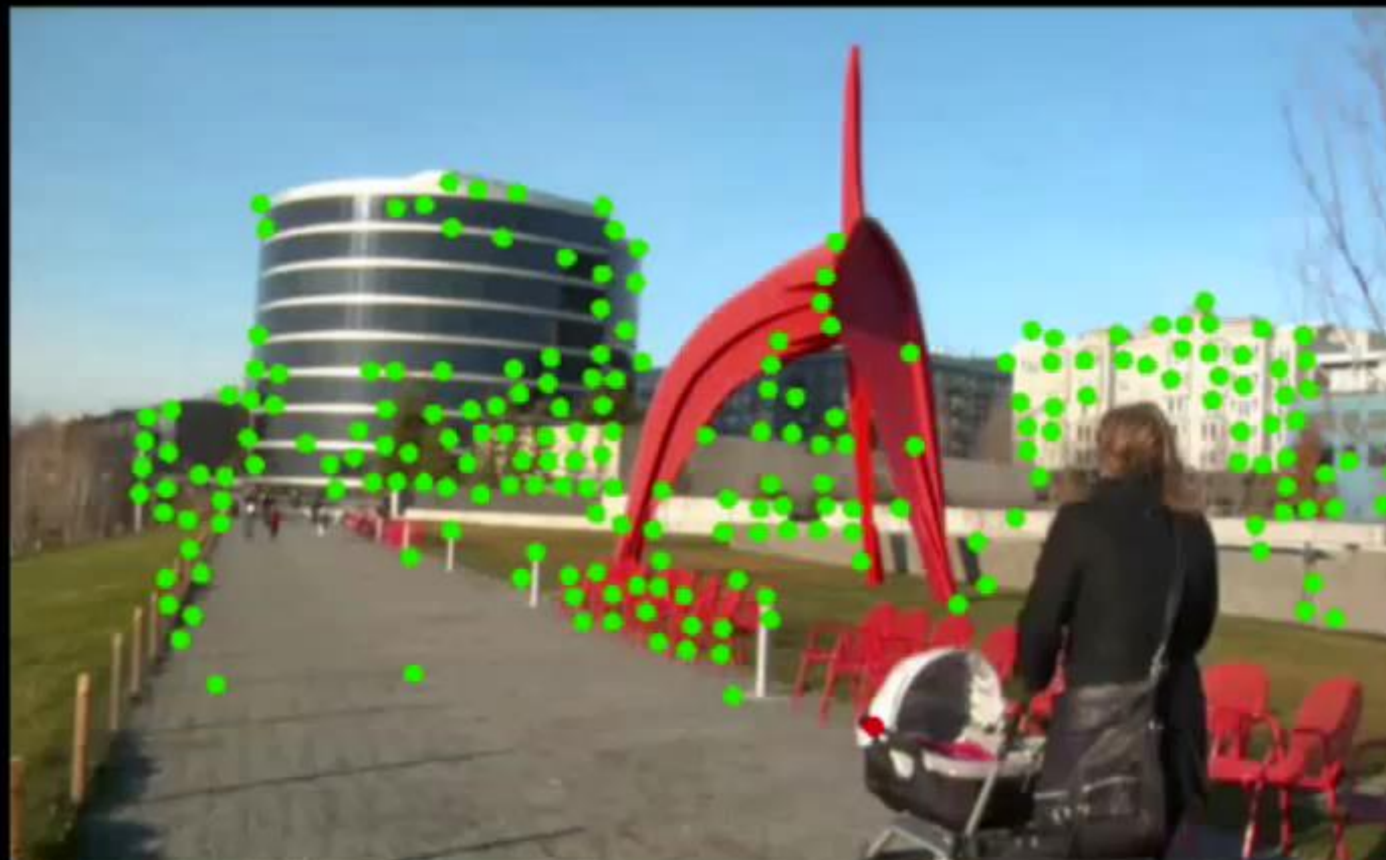
Result of fade-in/out the weight

# Question 1: How about Moving Objects?

- No reconstructed 3D points
- Warp follows surrounding background points

- Not correct! But…
  - ✓ Viewpoint shifts are small
  - ✓ Motion clouds the issue
  - ✓ Don’t notice exact occlusion relationships

# Question 2: How Is Novel View Synthesis?



Camera position

Output points

# Novel View Synthesis

# Novel View Synthesis

# Novel View Synthesis

# Results & Comparisons

INPUT

OUR OUTPUT

# Results & Comparisons

INPUT

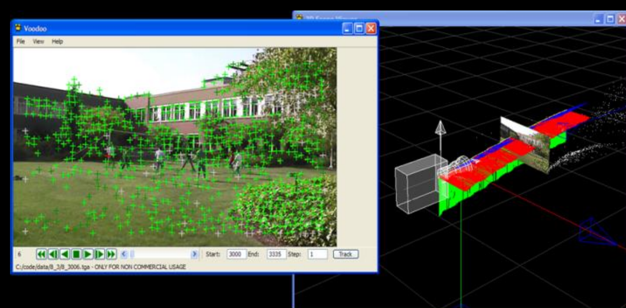
OUR OUTPUT

# Limitation

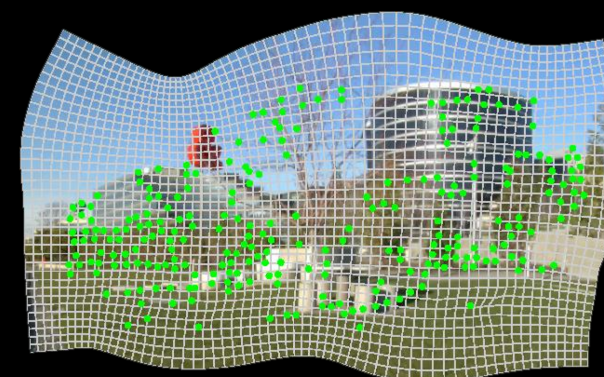
- Requires running structure-from-motion
  - Slow & memory-intensive
  - More brittle than simple point tracking
    - Need enough parallax
    - Difficult to deal with zooming
    - Vulnerable to artifacts in videos
  - Requires static regions to lock onto

# Our 3D Video Stabilization Method

Input → Trajectory Estimation √ → Motion Model Estimation → Motion Plan → Video Transform √ → Output

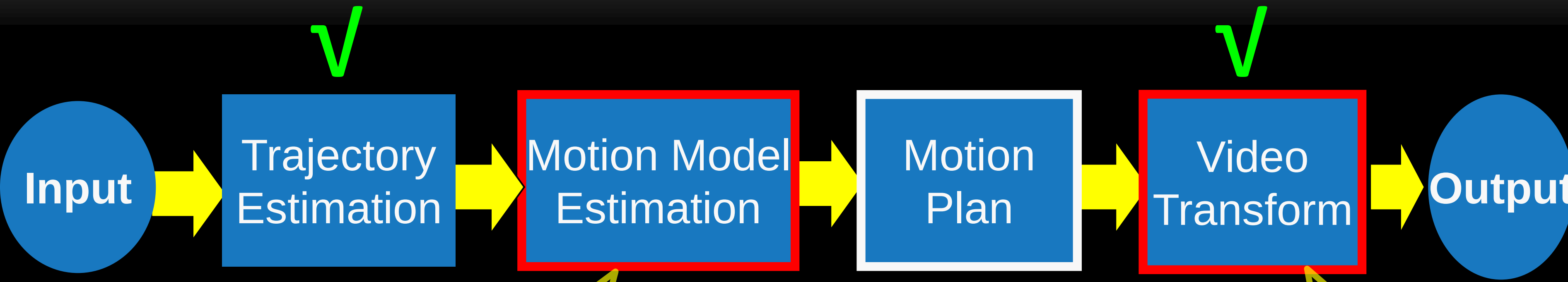


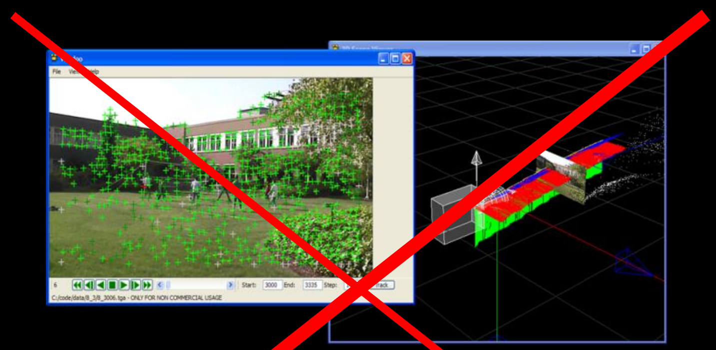
3D reconstruction via *structure from motion*



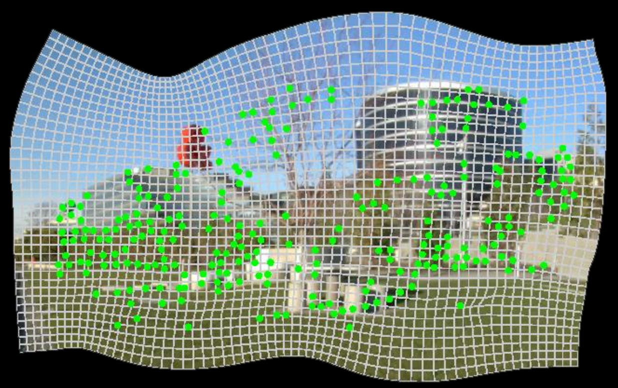
Content-preserving warp

# Video Stabilization Pipeline



~~3D reconstruction via *structure from motion*~~

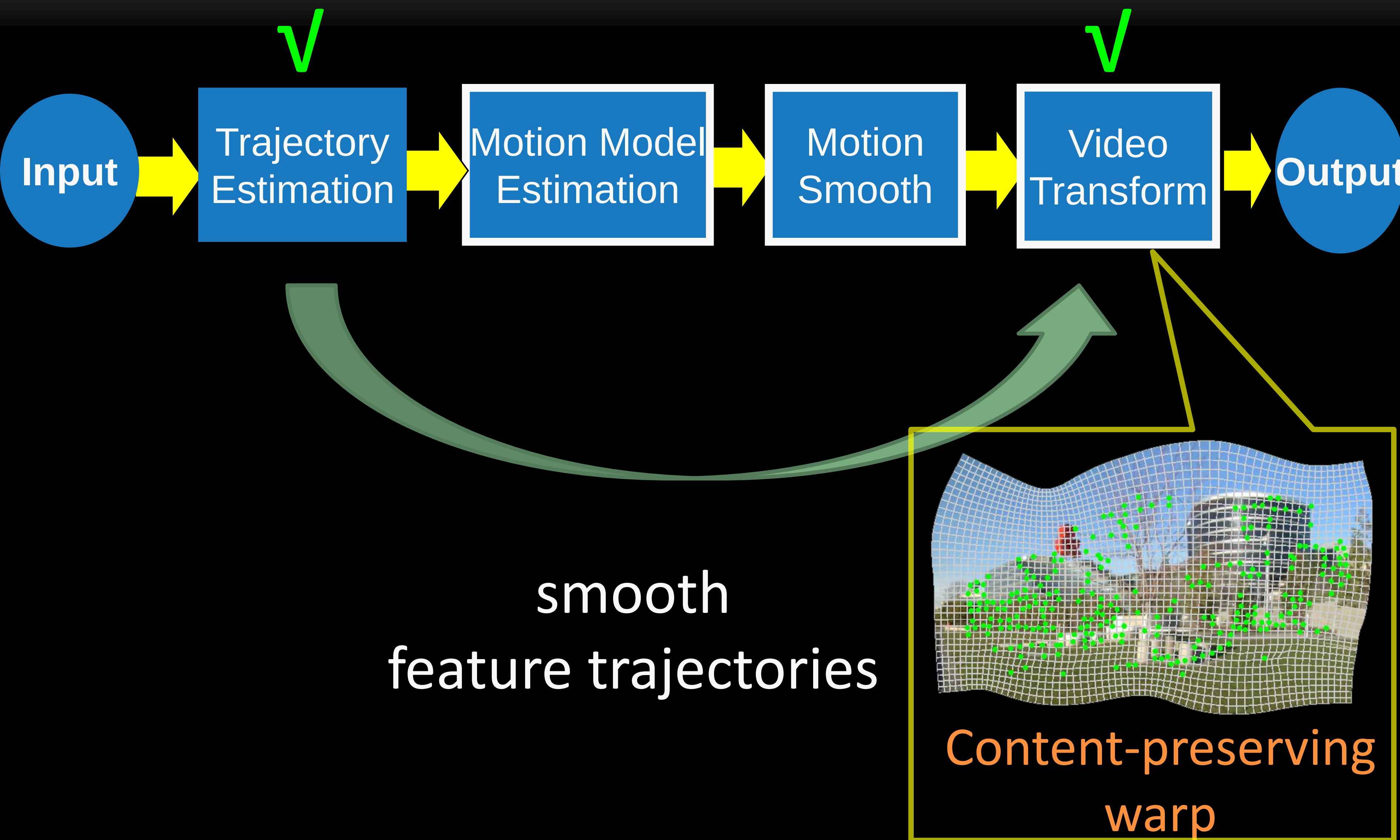


Content-preserving warp

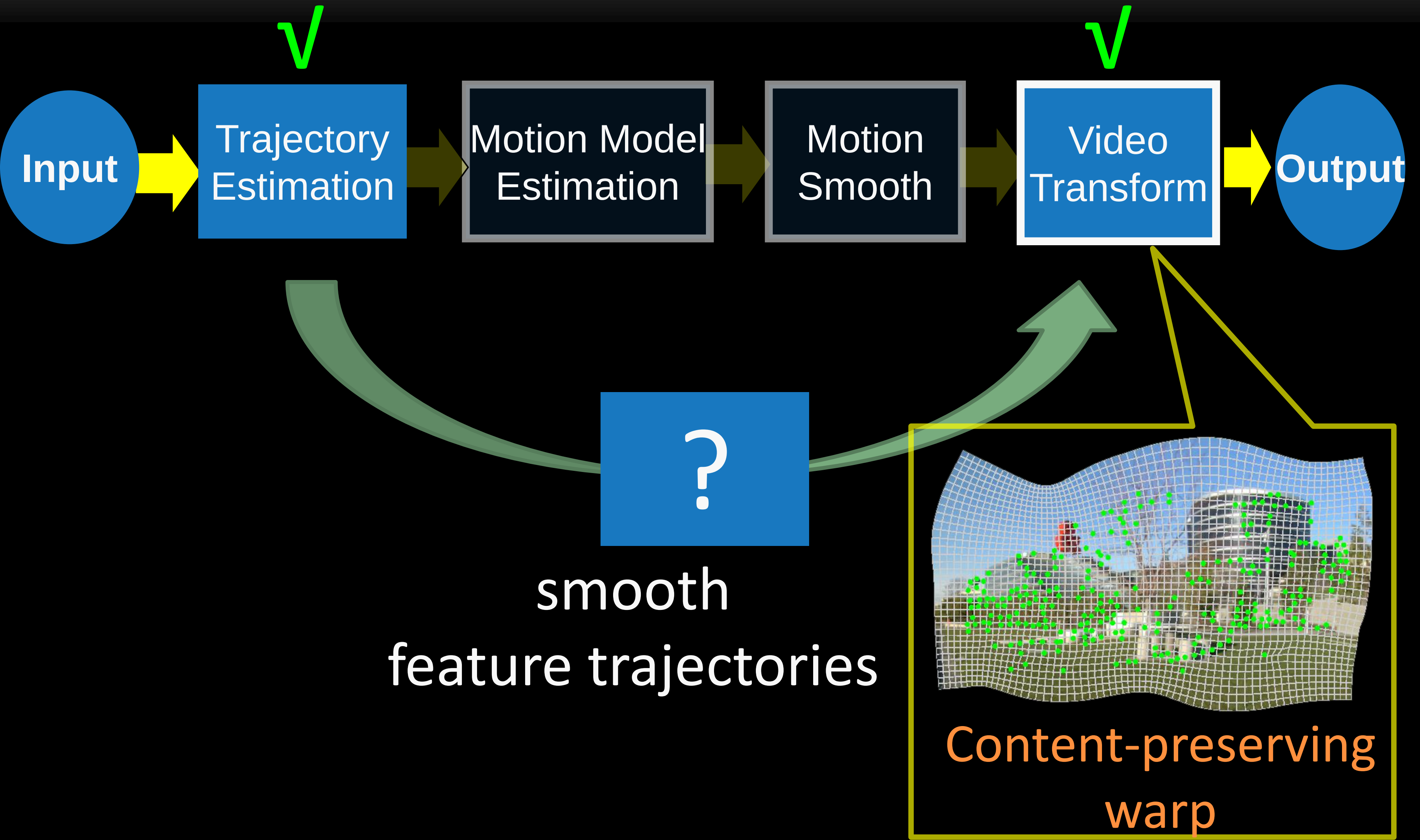
# Subspace video stabilization

[Liu et al., ACM Transactions on Graphics ‘11]

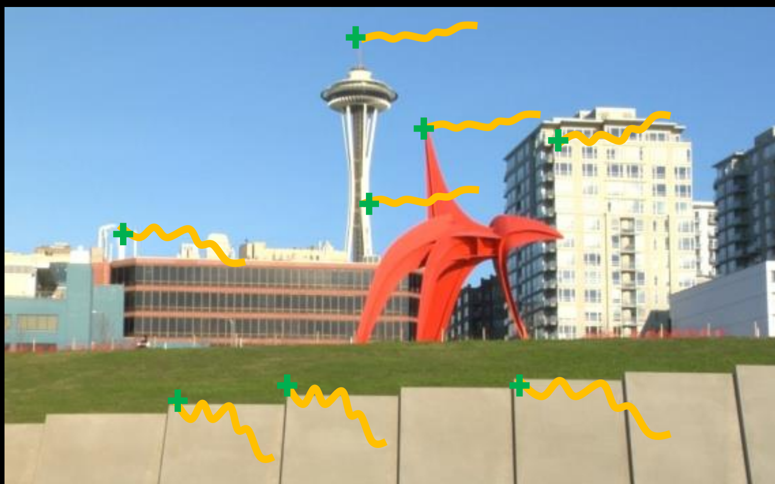
# Video Stabilization Pipeline



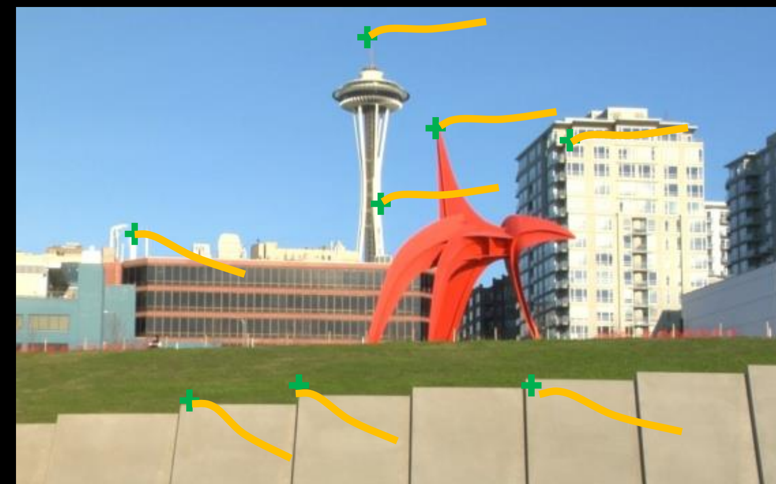
# Video Stabilization Pipeline



# Low-pass Filter Input Trajectories



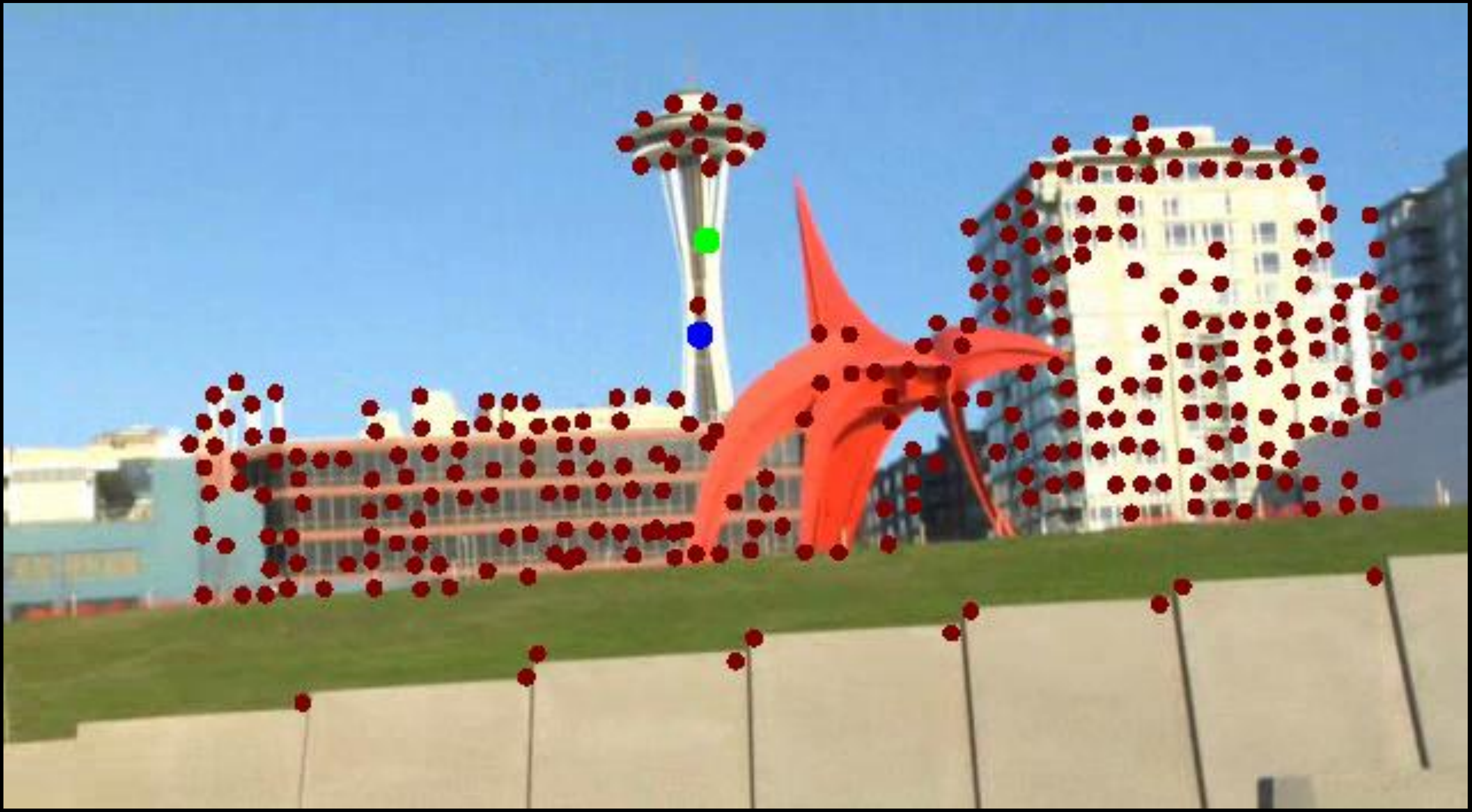
Filter



Input

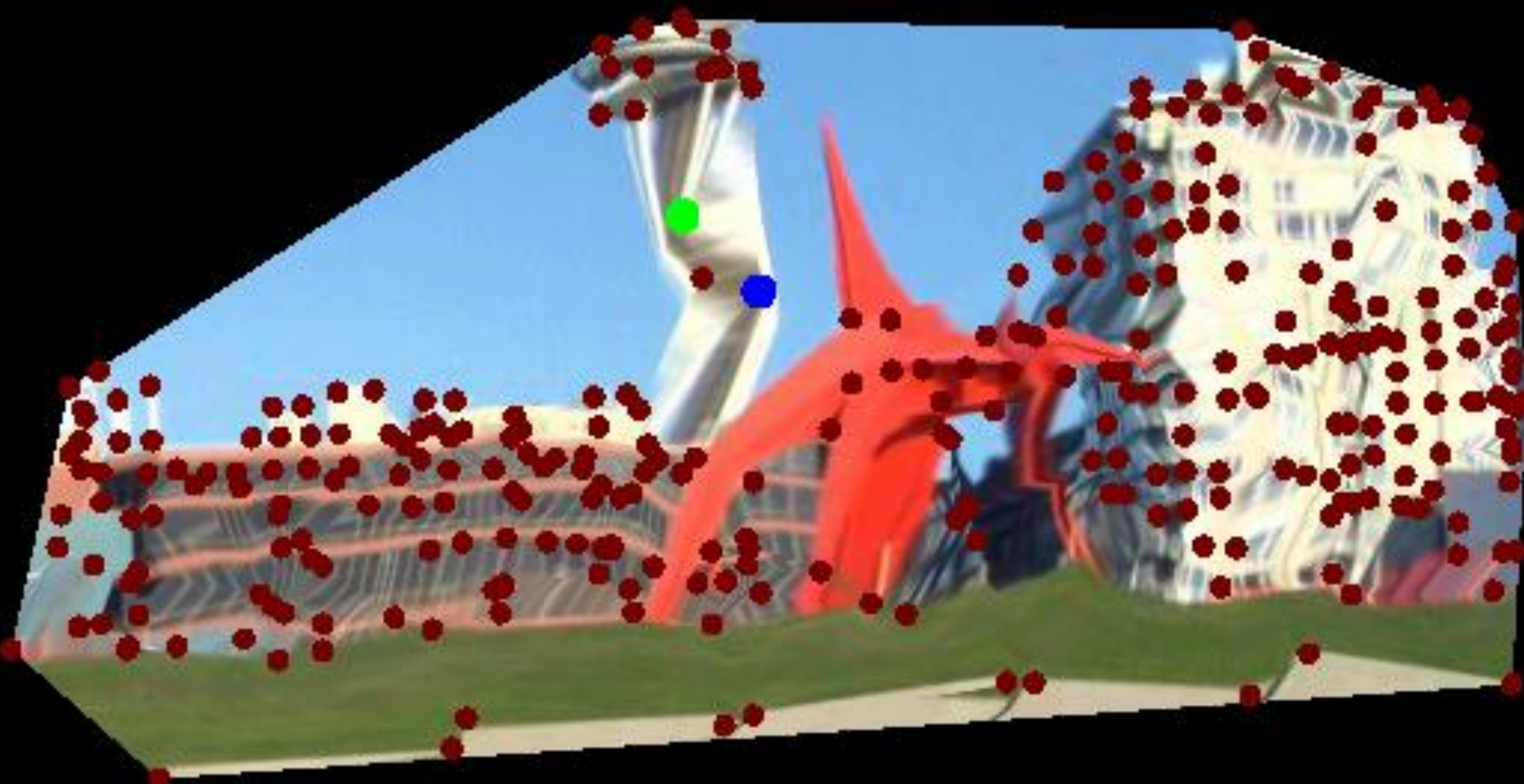
Output

# Low-pass Filter Input Trajectories



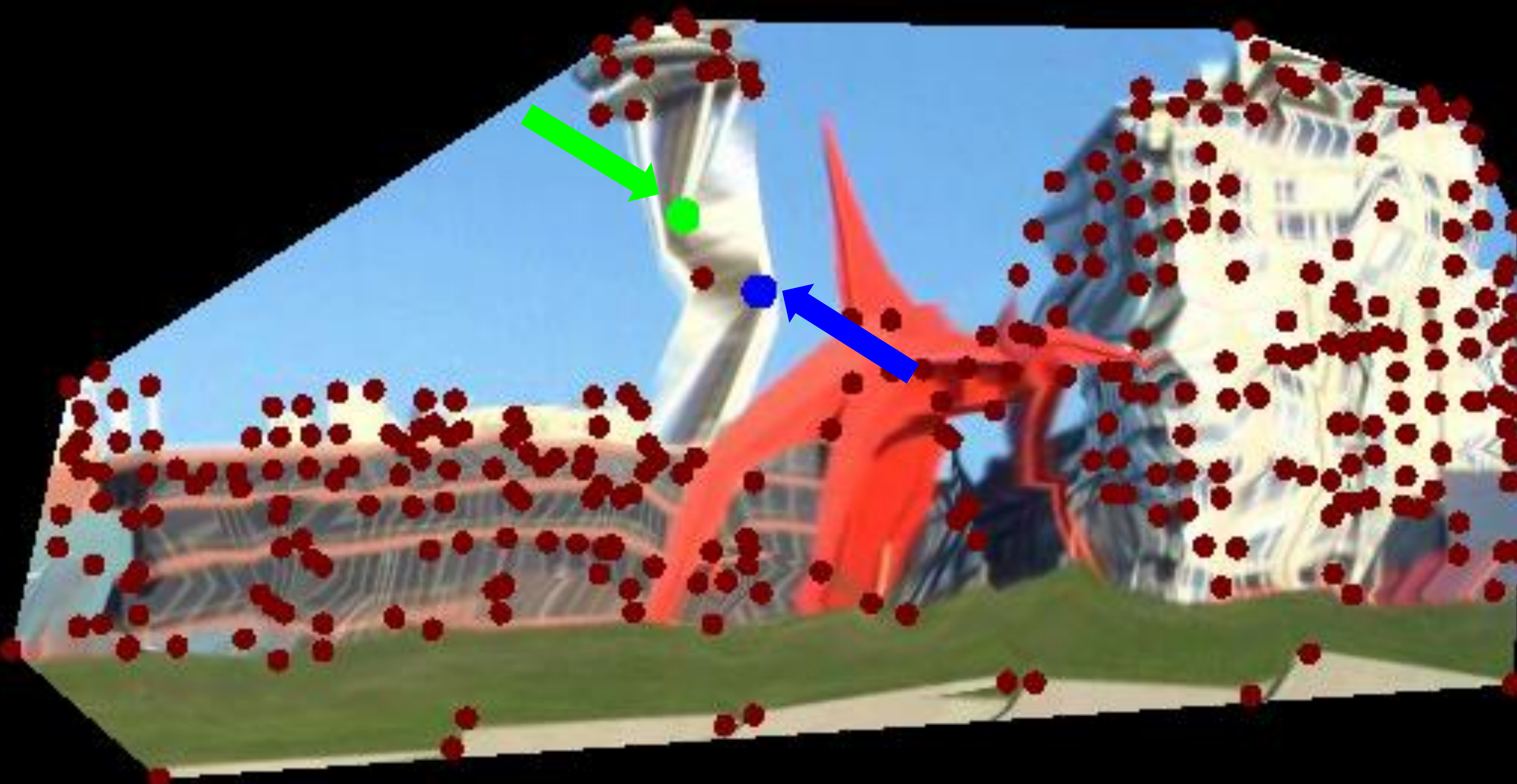
Input image

# Low-pass Filter Input Trajectories



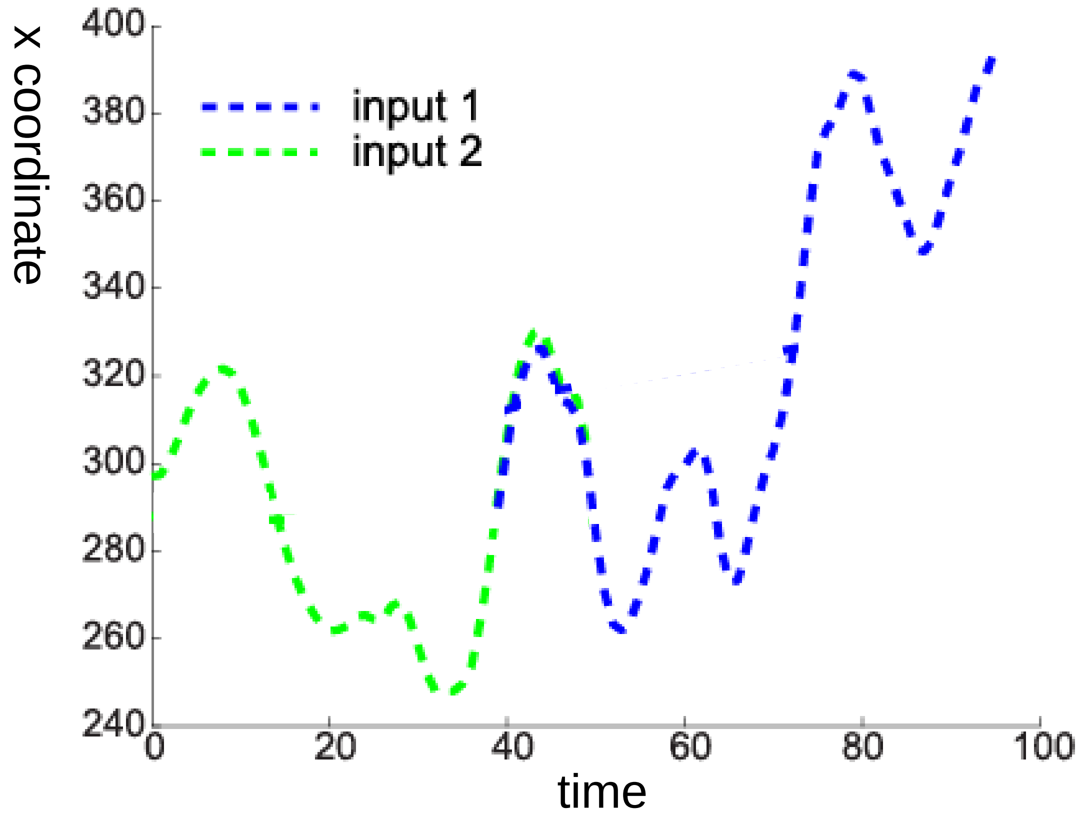
Result of filtering

# Low-pass Filter Input Trajectories

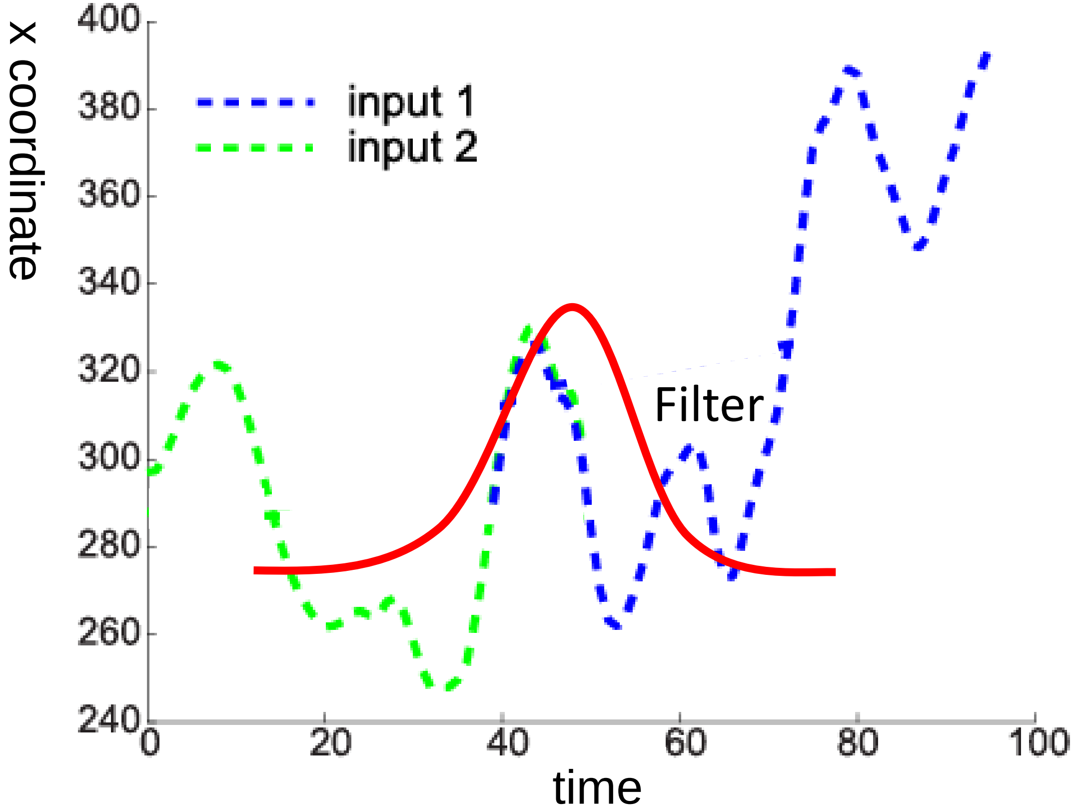


Result of filtering

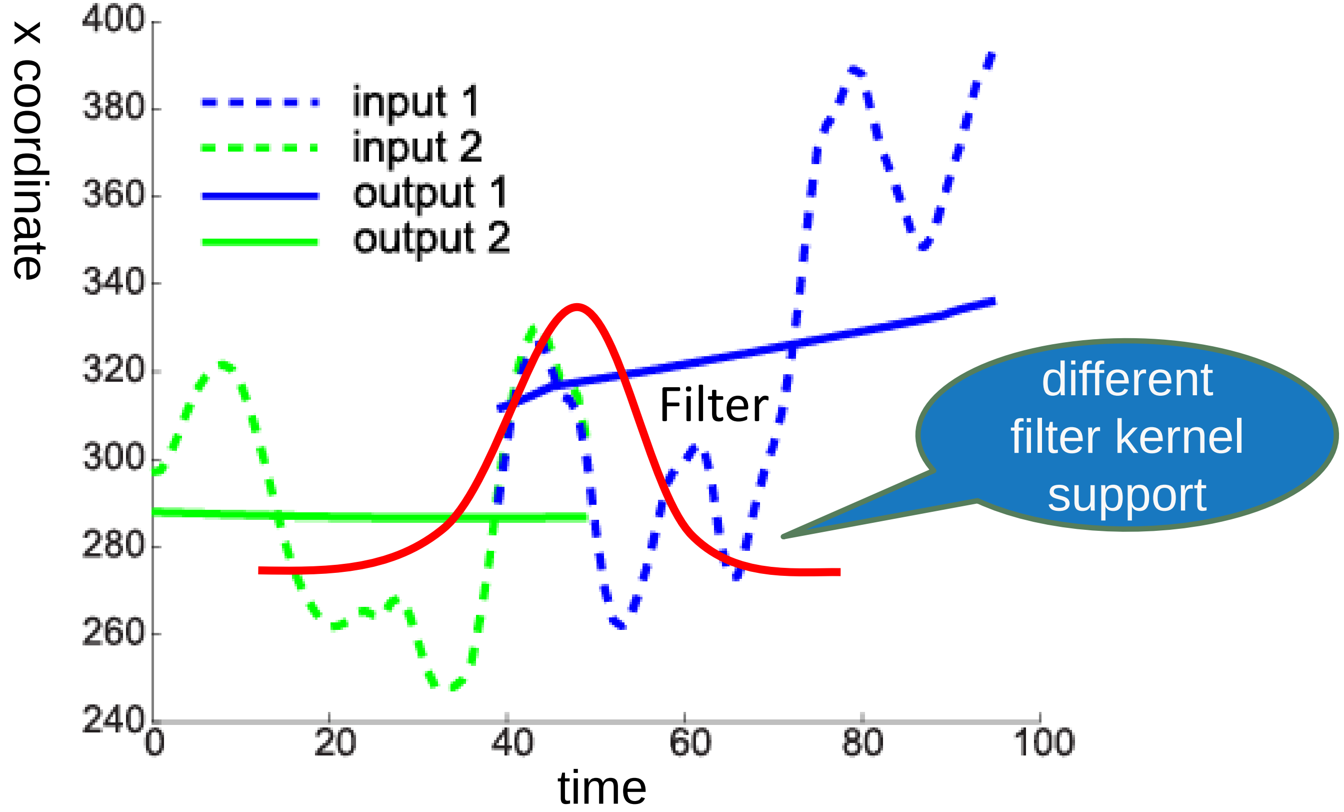
# Low-pass Filtering Trajectories



# Low-pass Filtering Trajectories



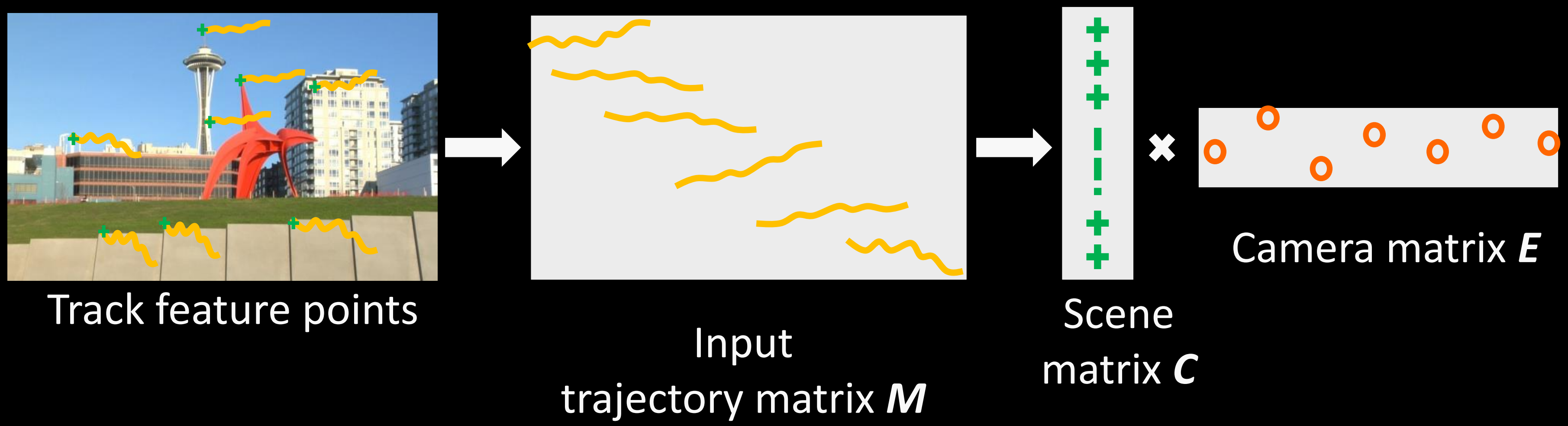
# Low-pass Filtering Trajectories



# Constraints

- Homography of 2D methods
  - Restrictive
  - Cannot model 3D feature motion
- Reconstructed 3D model of 3D methods
  - Difficult to estimate

# Affine Camera

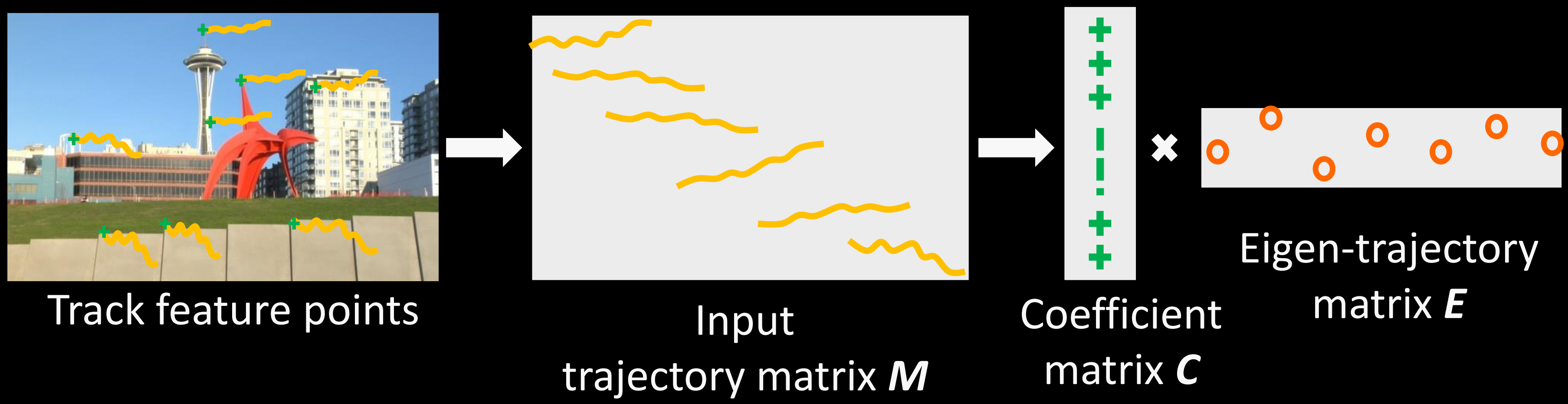


# Perspective Camera

The trajectory matrix of a rigid scene imaged by a moving camera over a short period of time should approximately lie in a low-dimensional subspace.

[Tomasi and Kanade 1992, Irani 2002]

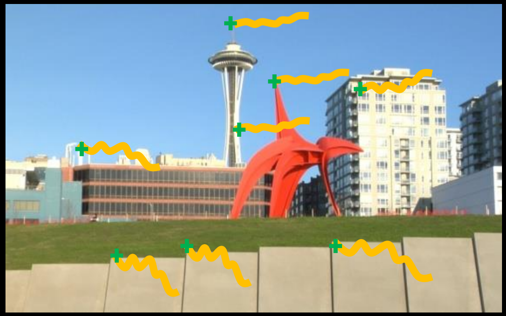
# Perspective Camera




The trajectory matrix of a rigid scene imaged by a moving camera over a short period of time should approximately lie in a low-dimensional subspace.

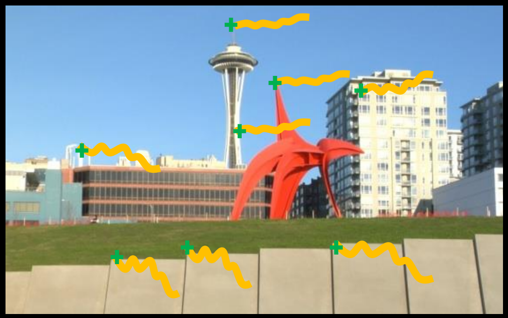
[Tomasi and Kanade 1992, Irani 2002]

# Subspace Stabilization Overview

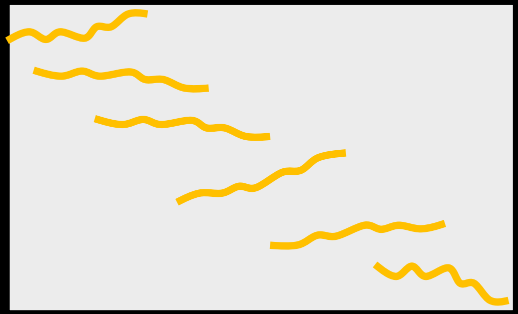


Track feature points

# Subspace Stabilization Overview

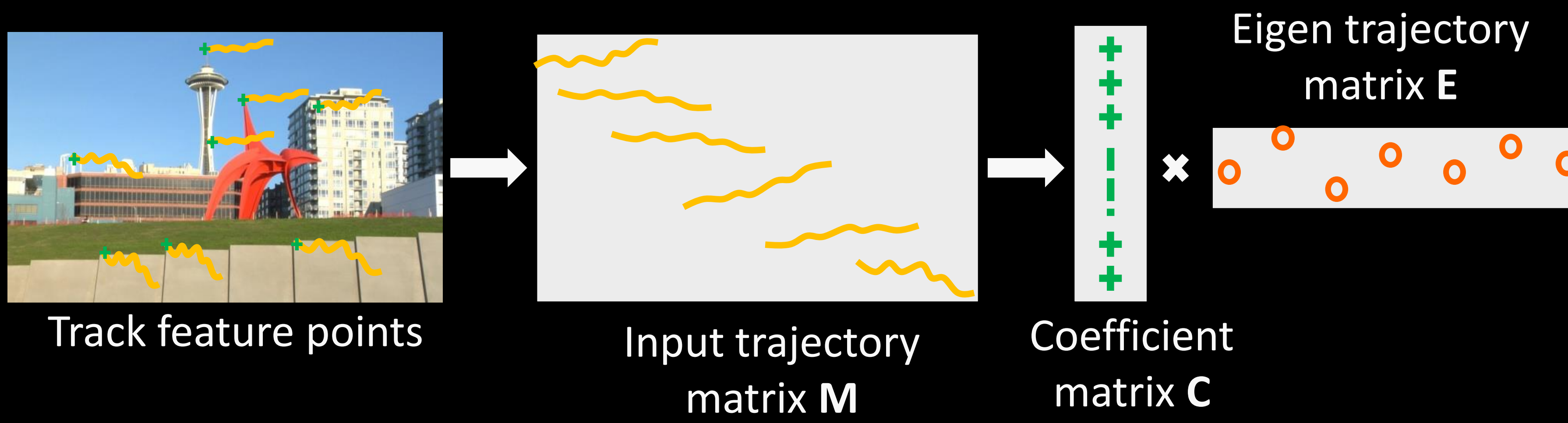


Track feature points

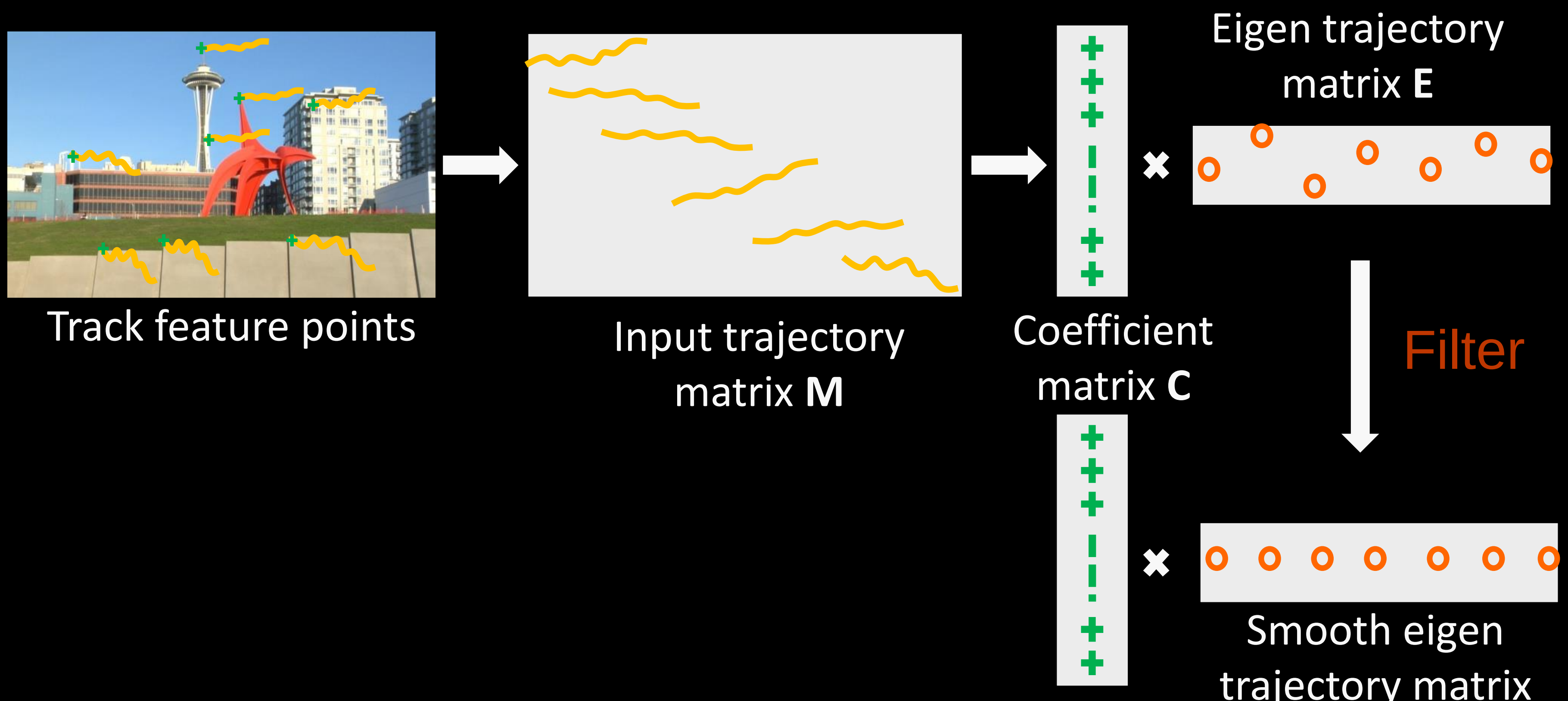


Input trajectory matrix **M**

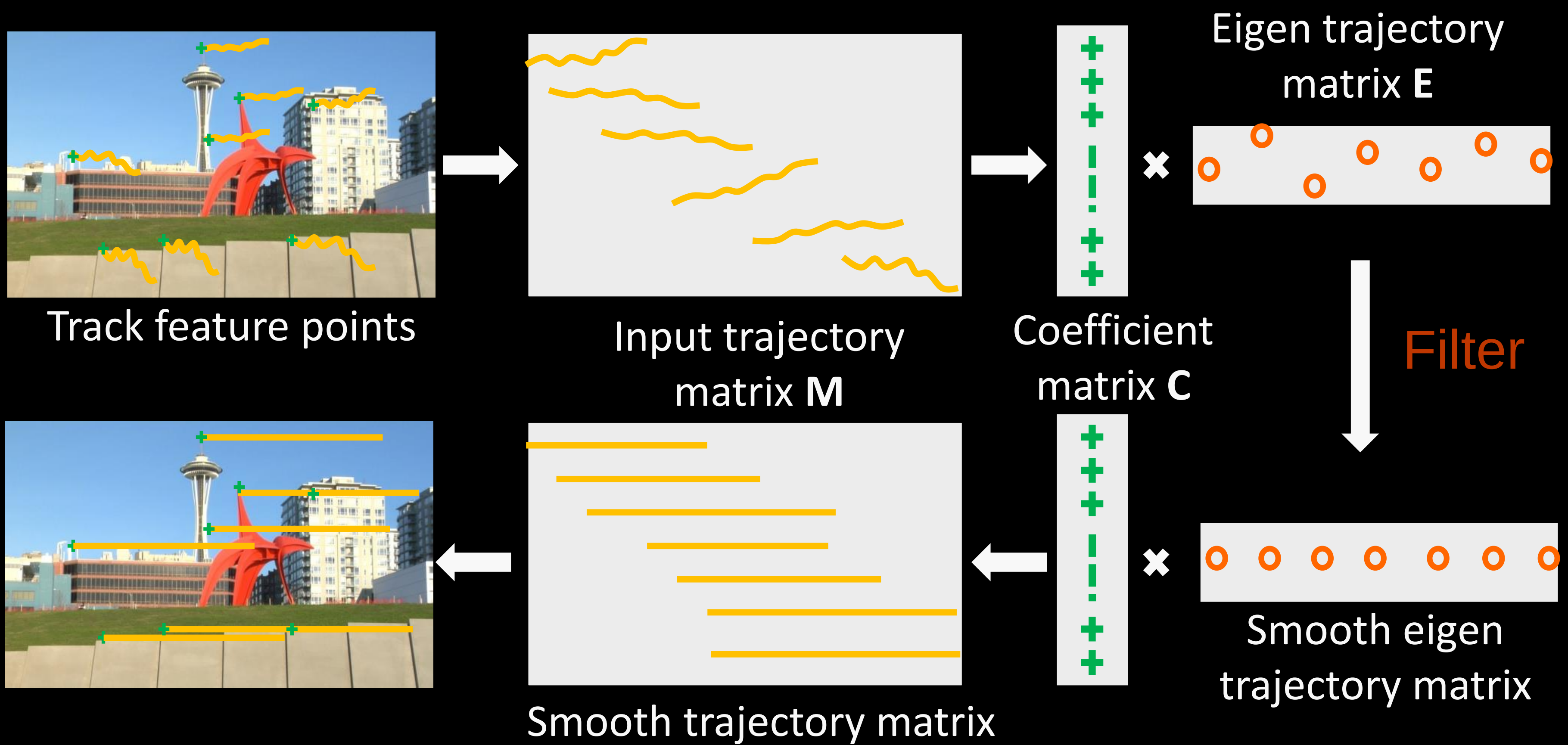
# Subspace Stabilization Overview



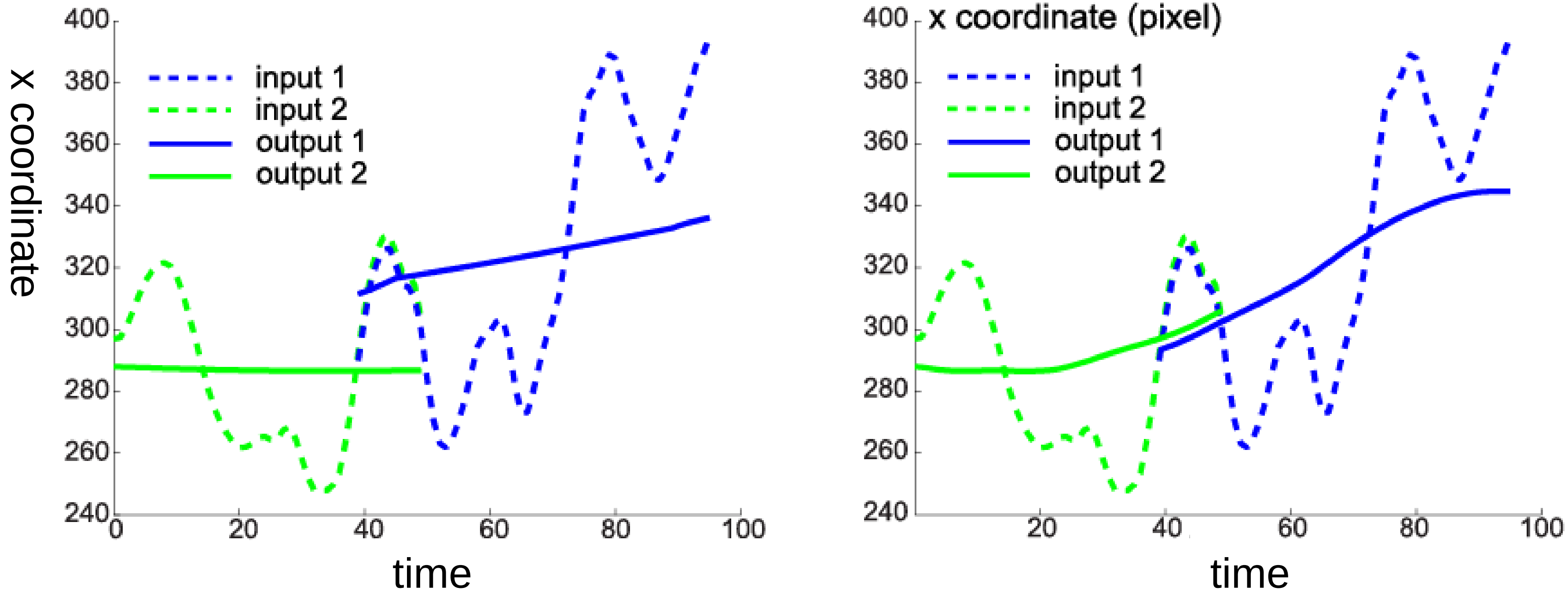
# Subspace Stabilization Overview



# Subspace Stabilization Overview



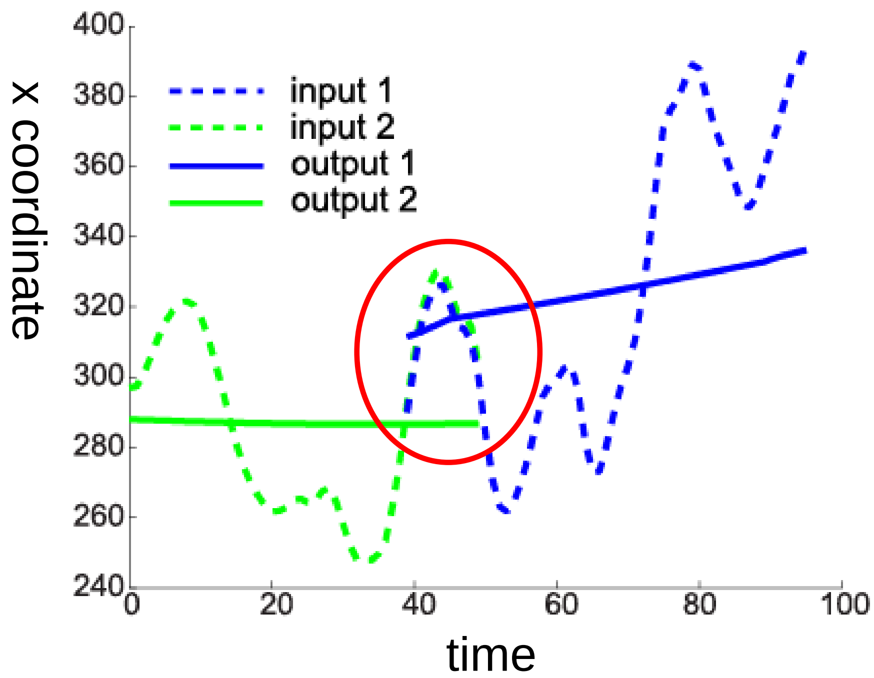
# Low-pass Filtering Trajectories



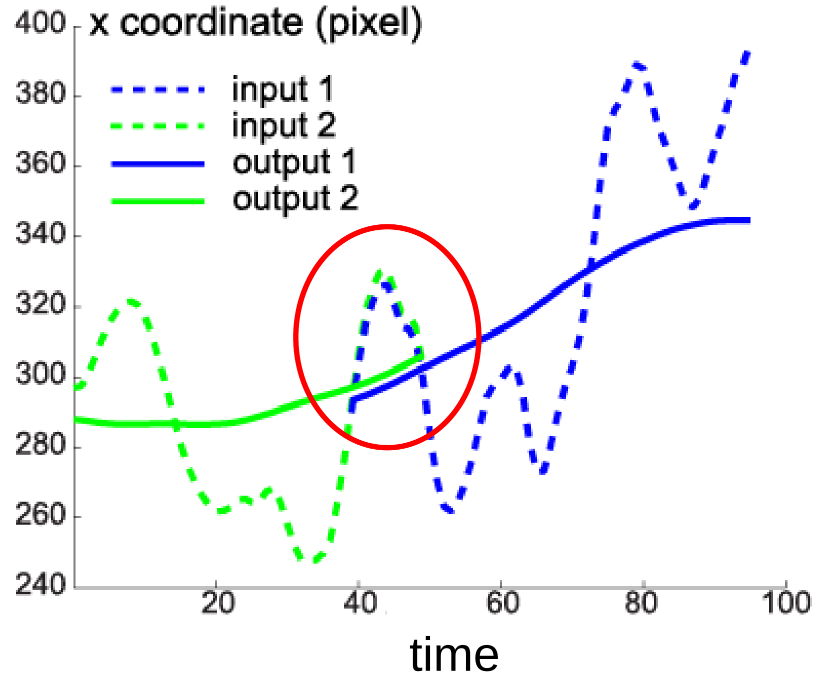

Naïve filtering

Subspace filtering

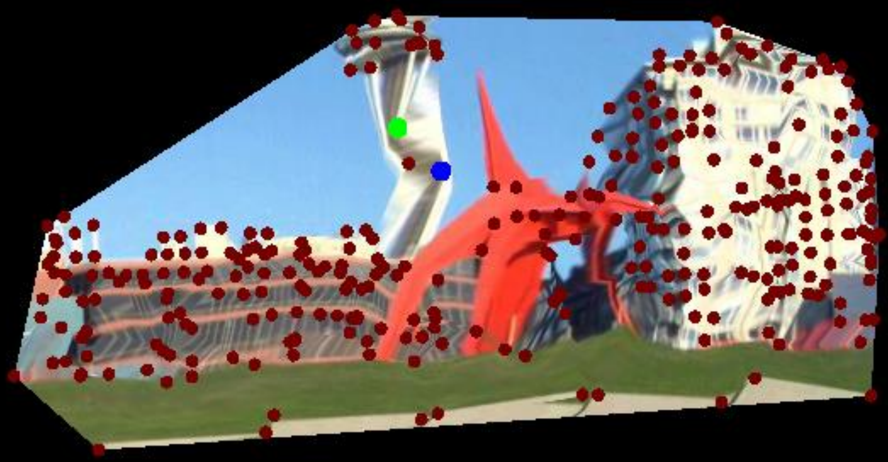
# Low-pass Filtering Trajectories



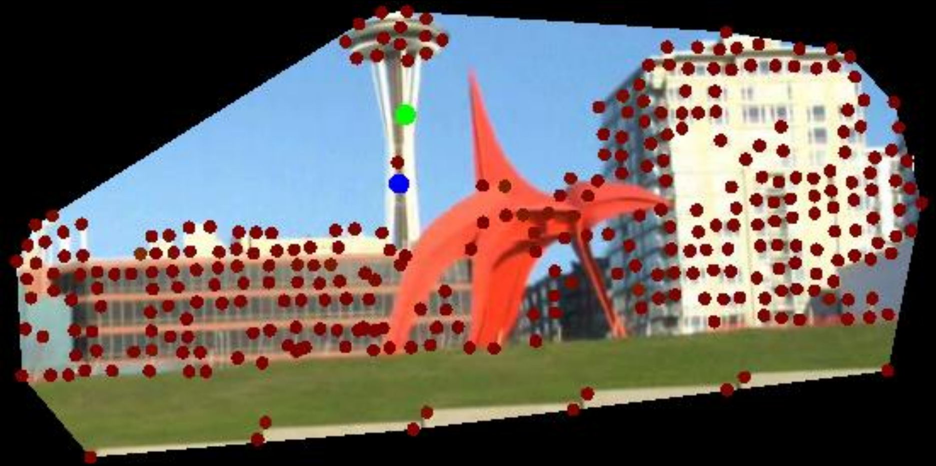

Naïve filtering




Subspace filtering

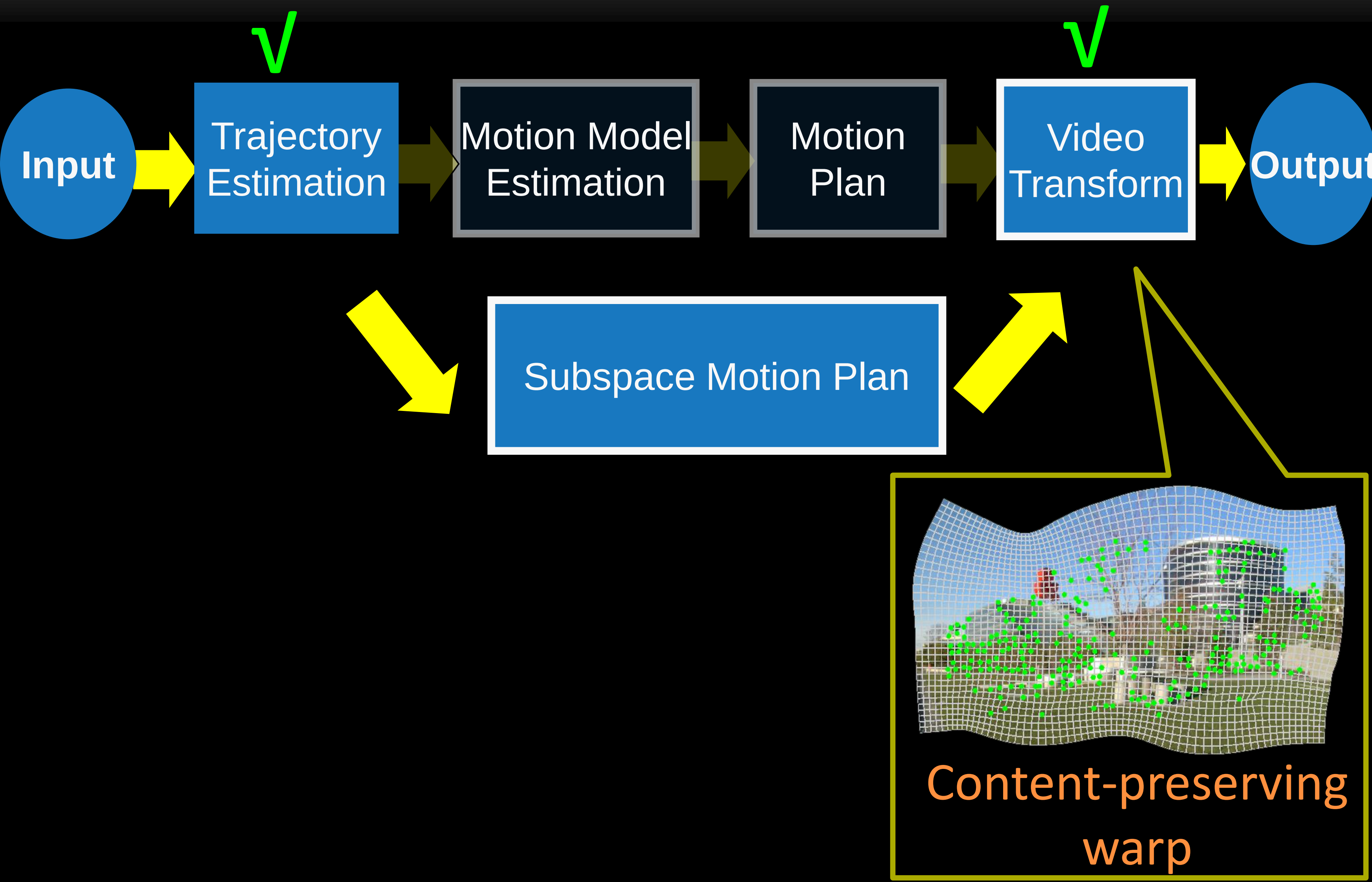


Naïve filtering

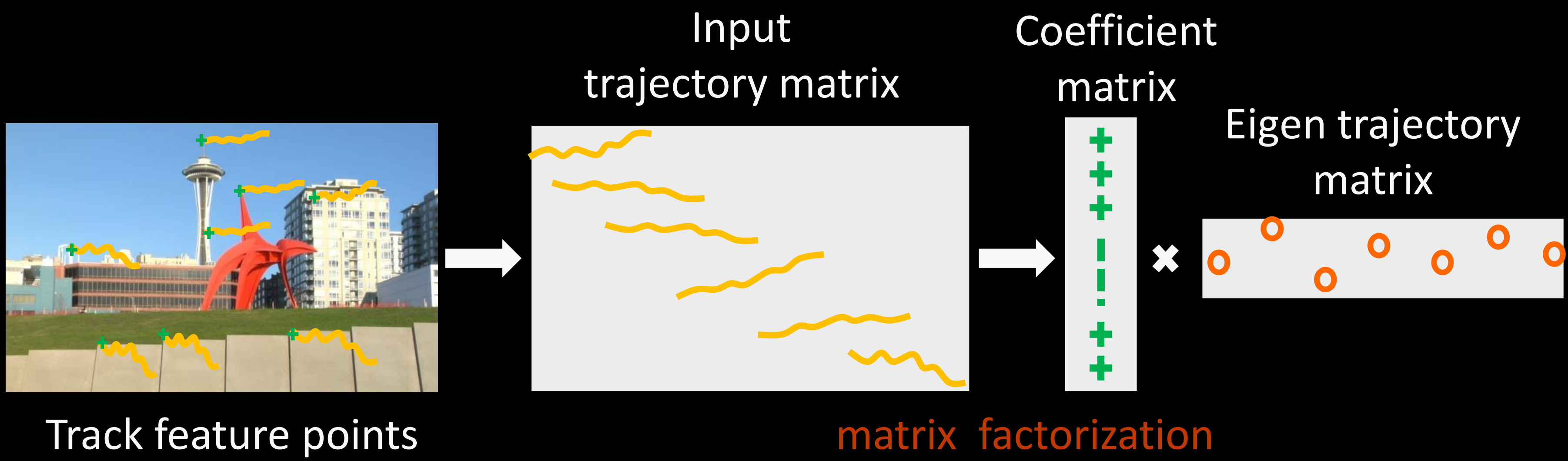


Subspace filtering

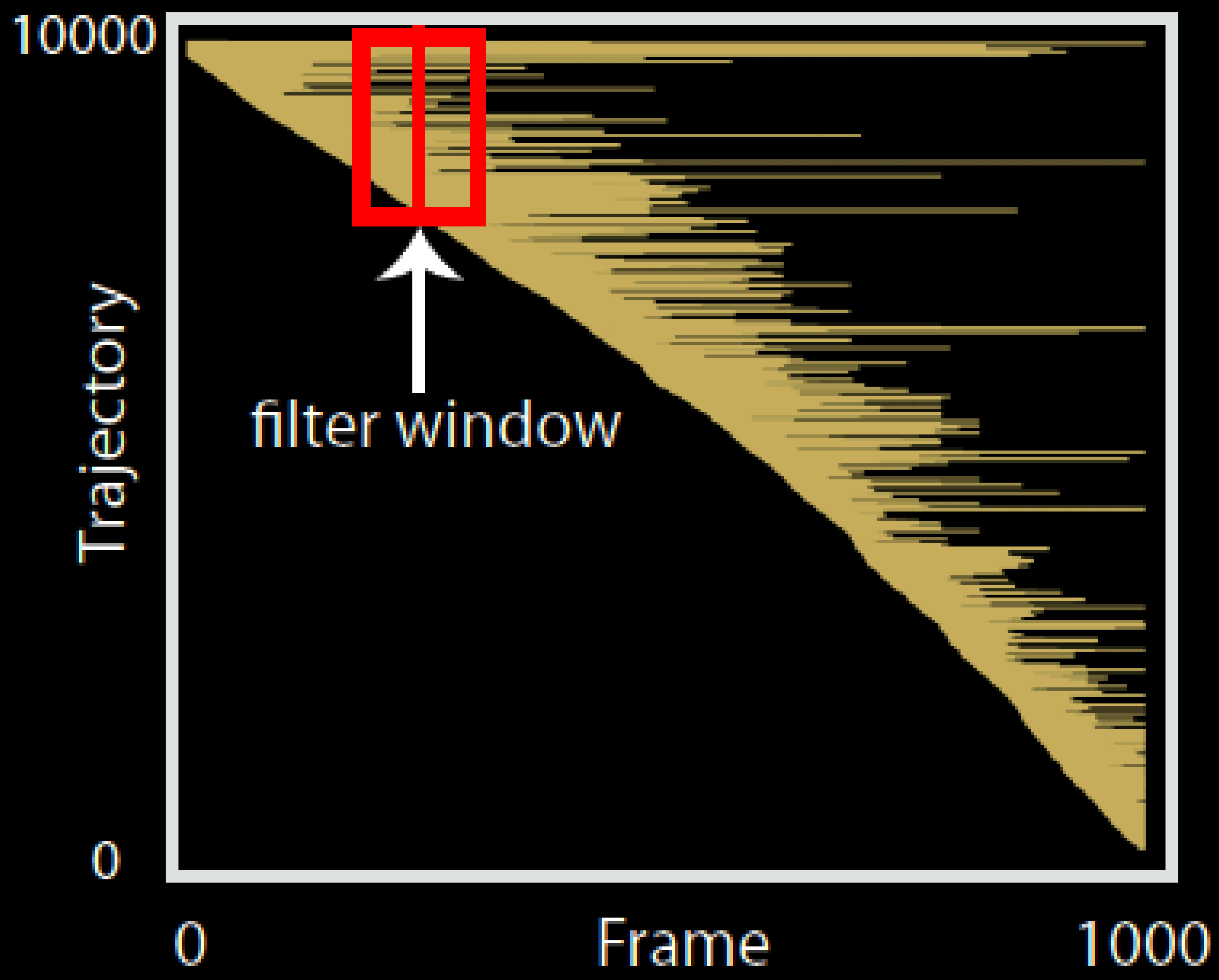
# Video Stabilization Pipeline



# Matrix Factorization



# Incomplete Trajectory Matrix



# Incomplete matrix factorization

- Iterative methods [Buchanan and Fitzgibbon 05, Chen 2008]
  - Accurate
  - Time-consuming and not streamable
    - Difficult to handle long videos.
- Moving factorization
  - A variation of incremental factorization methods
  - Greedy method and less accurate
    - But, good enough for stabilization
  - Efficient and streamable

# Factorization accuracy

- Testing data: 70 videos
  - Resized to 640 x 360
- Moving factorization
  - Error: 0*.08 to 0.26 pixels*
- Iterative factorization
  - ¼ of moving factorization errors on average

# Result

Input

Our result

# Result: A Long Video

Input

Our result

# Comparison with Our 3D Method

Input

3D

Our result

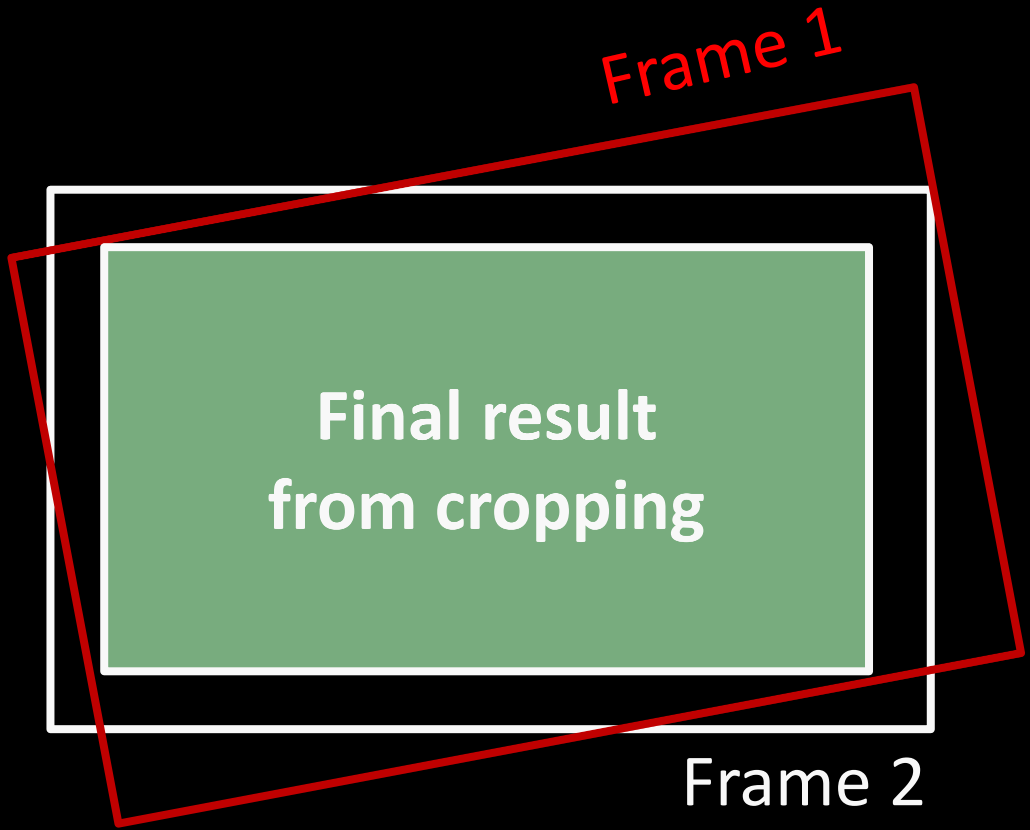
# Demo

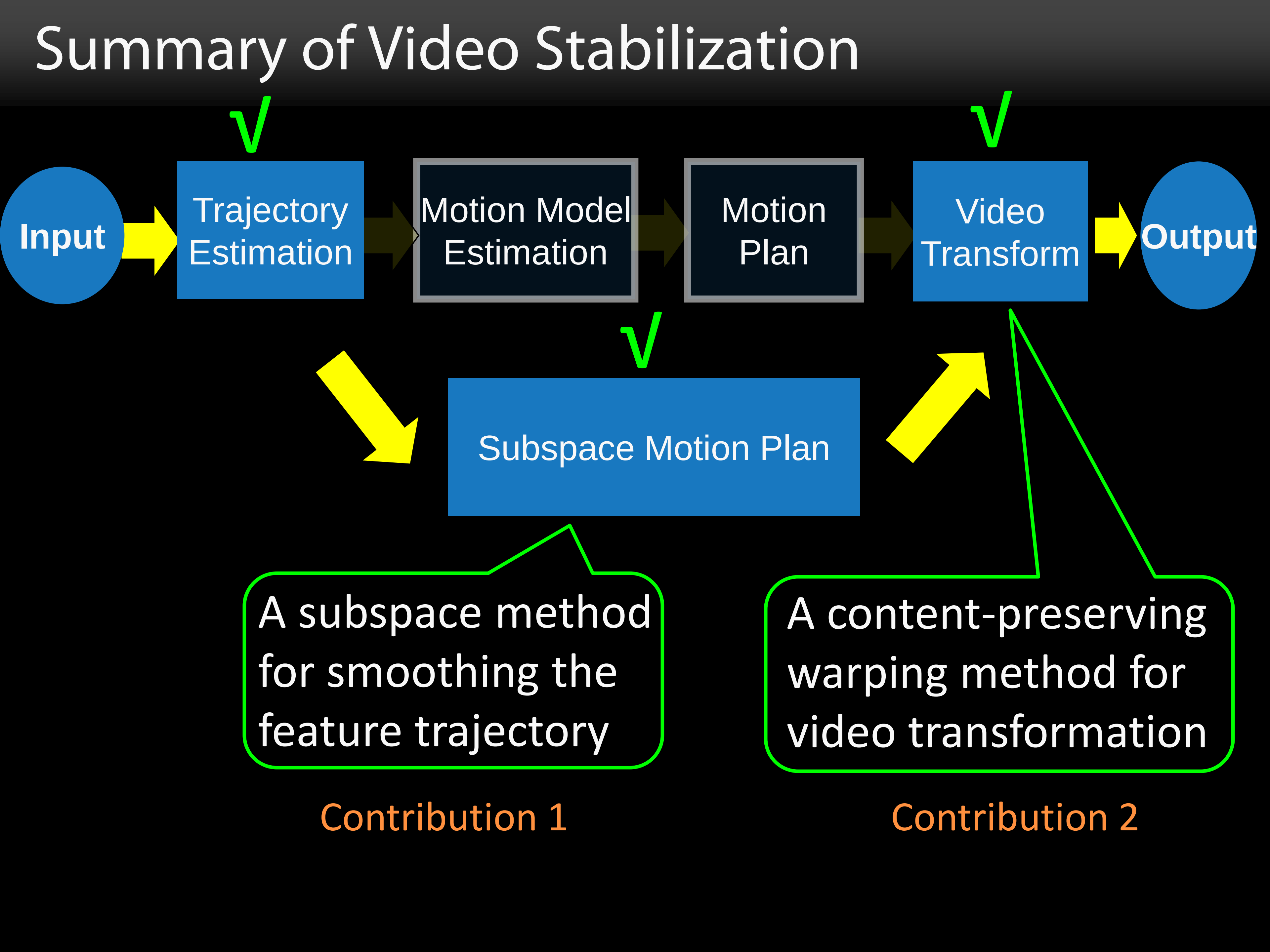
Input

Our result

# Limitations

- Aggressive stabilization leads to aggressive cropping
- Need reasonably long feature trajectories




Summary of Video Stabilization
√
√
√
Input
Trajectory Estimation
Motion Model Estimation
Motion Plan
Video Transform
Output
Subspace Motion Plan
A subspace method for smoothing the feature trajectory
A content-preserving warping method for video transformation
Contribution 1
Contribution 2

# Impact

✓ Selected as one of the five Top Videos of 2009 by *New Scientist*

✓ *Warp Stabilizer* in After Effects CS5.5 and Premiere CS 6.0 is largely based on our research



# Student paper presentation

## Video SnapCut: Robust Video Object Cutout Using Localized Classifiers

X. Bai, J. Wang, D. Simons, G. Sapiro.
SIGGRAPH 2009

**Presenter:** Wiemholt, Cody

# Student paper presentation

## A Global Sampling Method for Alpha Matting

K. He, C. Rhemann, C. Rother, X. Tang, and J. Sun
CVPR 2011

**Presenter:** Zwovic, Kitt

# Next Time

- Video Stabilization III
- Stereoscopy Photography
- Student paper presentation
  - 05/19: Filgas, Ryan
    - A Closed Form Solution to Natural Image Matting
      A. Levin, D. Lischinski, and Y. Weiss
      CVPR 2006