Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Digital Video

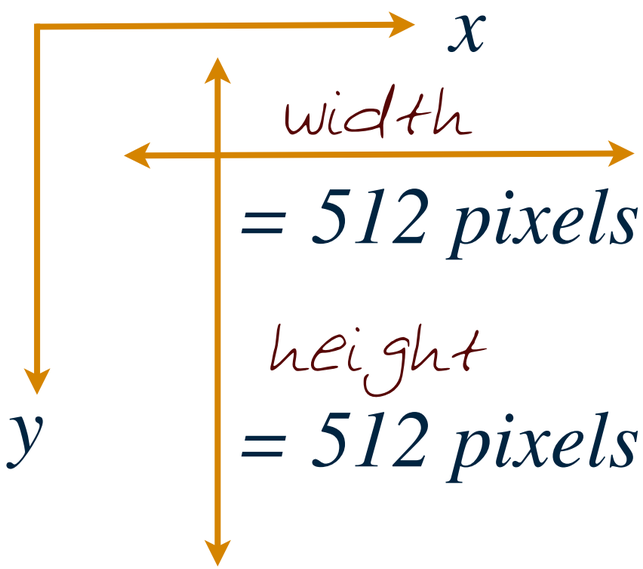
* Video is basically just a stack of images in Time

Lesson Objectives

1. Relationship between Images and Videos
2. Persistence of vision in playing (and capturing) Videos
3. Extend filtering and processing of Images to Videos
4. Tracking points in Videos

Recall: A Digital Image

Georgia Tech's Mascot Buzz, in Black and White

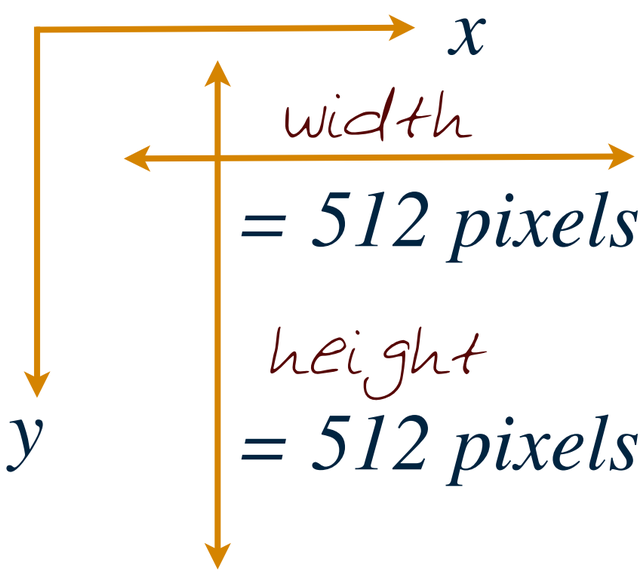


512x512 pixels
= 262,144 pixels
= .26 MP image

* "Digital" Image:
  * numeric representation in two-dimensions (x and y)
  * referred to as $I(x,y)$ in continuous function form, $I(i,j)$ in discrete

Recall: A Digital Image

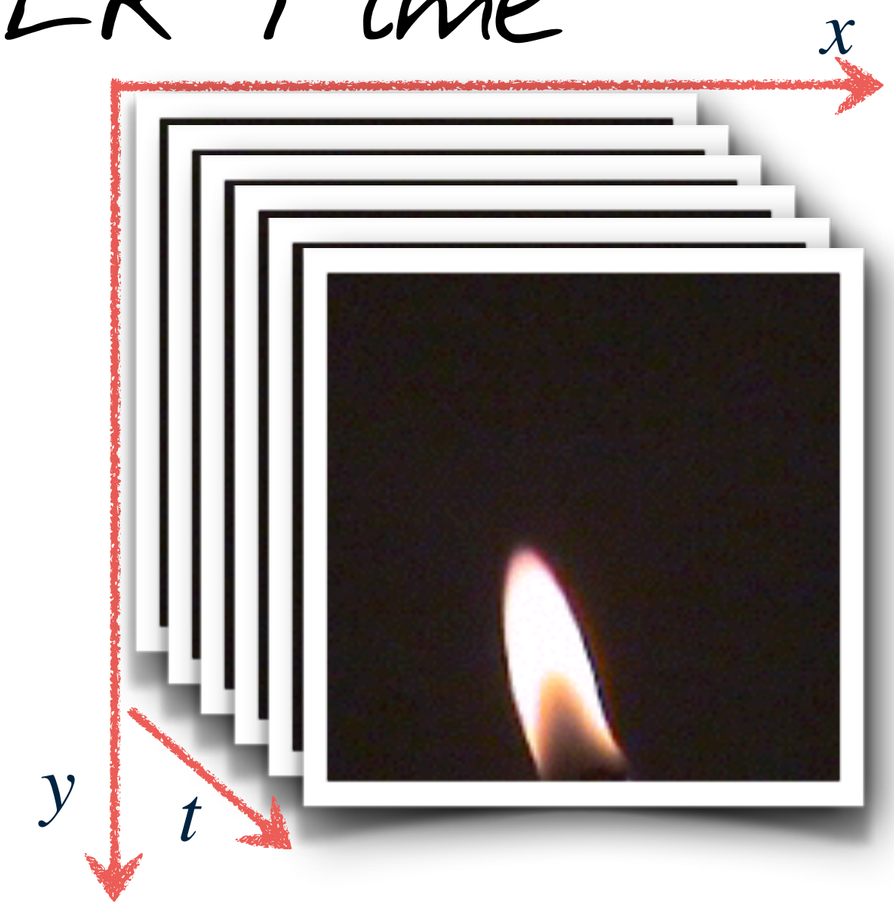
Georgia Tech's Mascot Buzz, in Black and White



512x512 pixels
= 262,144 pixels
= .26 MP image

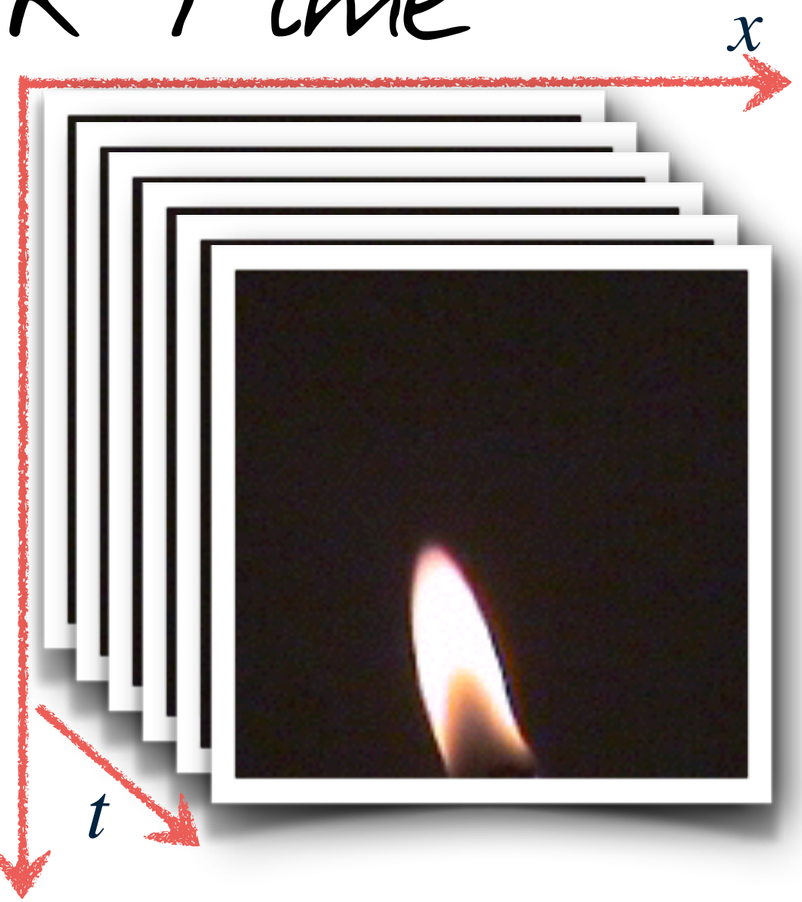
* Image Resolution:
  * expressed as representation of Width and Height of the image
* Each pixel (picture element) contains light intensities for each value of x and y of $I(x,y)$

Video: Images OVER Time



* "Digital" Video:
  * numeric representation in two-dimensions (x and y), stacked in time, t
  * referred to as $I(x,y,t)$ in continuous function form, $I(i,j,t)$ in discrete

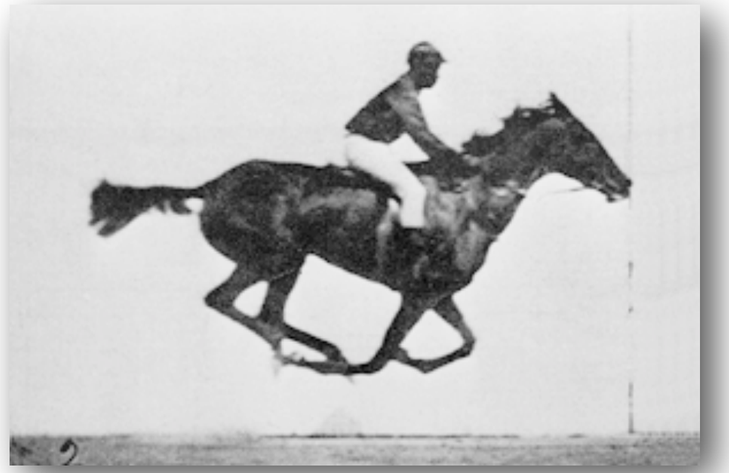
Video: Images OVER Time



* Video Resolution:
  * expressed as representation of Width and Height of the image
  * Usually in aspect ratios of **4x3**, **16x9**, etc
* File formats: Include images, frame-rates, and codec/wrappers

Persistence of Vision

* If image frames are captured and played back (refreshed) at a rate faster **1/24th** of a second
* We see flicker-free appearance of motion



en.wikipedia.org/wiki/File:Muybridge_race_horse_animated.gif Pictured in 1887, Animated in 2006

en.wikipedia.org/wiki/File:Marey_-_birds.jpg

Persistence of Vision

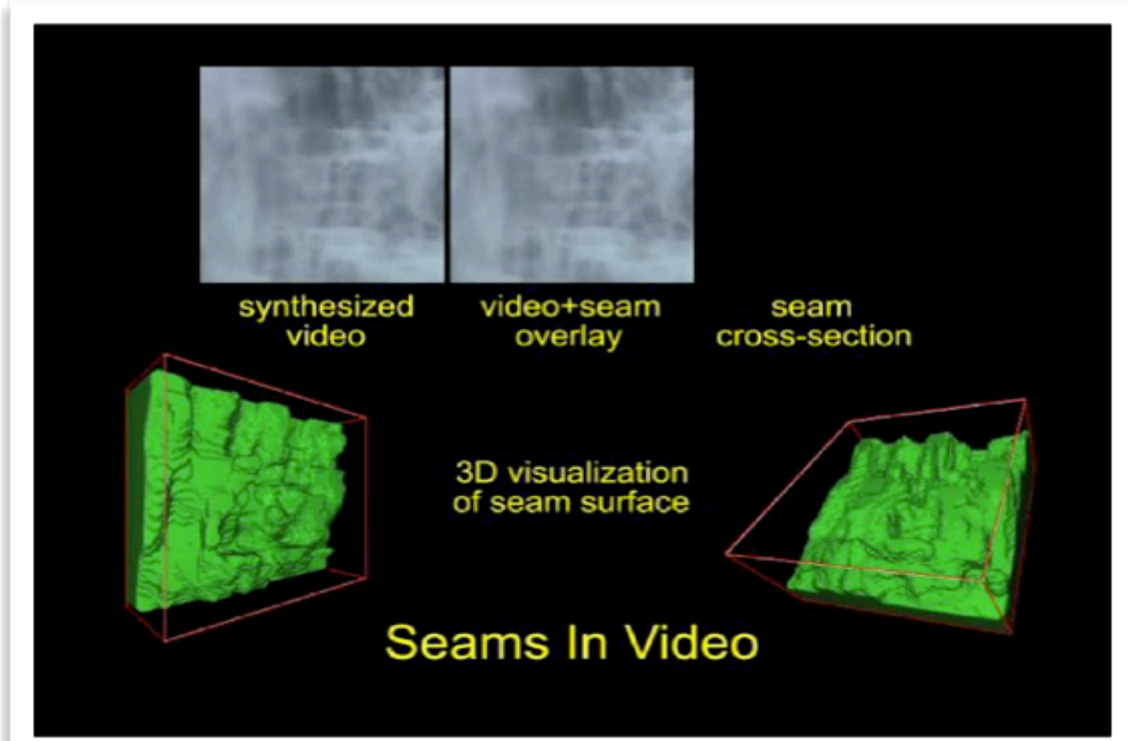
* Foundational observation of why we perceive video
* Rationale behind the invention of video cameras
* Muybridge (1830-1904) used stop-action photographs to study animal motion
* Marey (1830-1904) developed Chronophotographe to capture motion

en.wikipedia.org/wiki/File:Muybridge_race_horse_animated.gif Pictured in 1887, Animated in 2006

en.wikipedia.org/wiki/File:Marey_-_birds.jpg

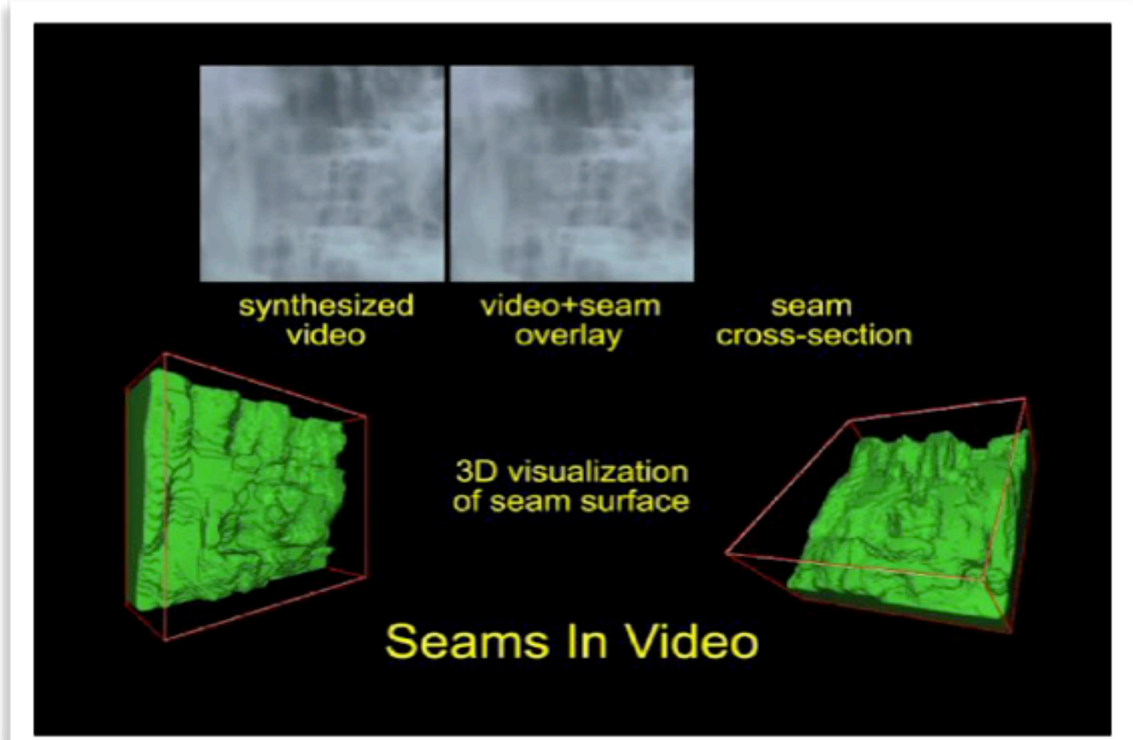
Processing/Filtering Video

* Same as with images, just over a video volume
* Can filter in 3D
  * (x, y, t)
* Motion information is used in video compression



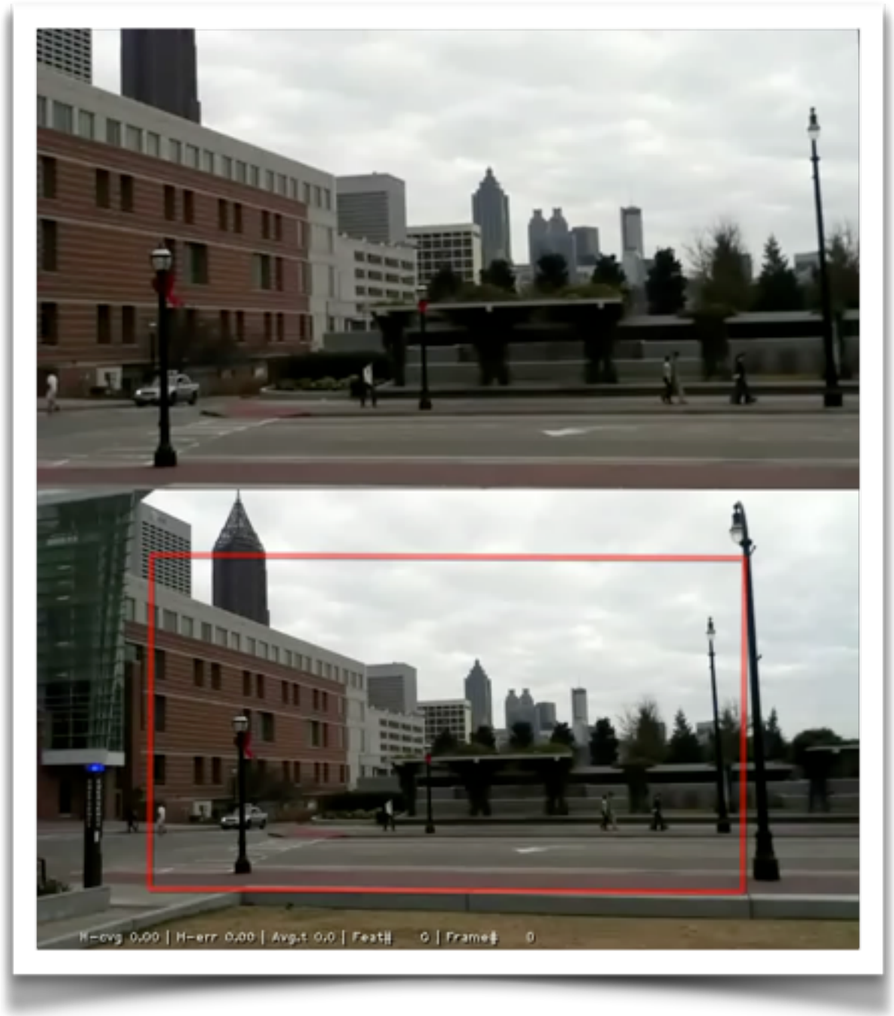
Processing/Filtering Video

* Same concepts of change detection as in xy- space
  * apply to xt- and yt- space.
* If all pixels from one frame to another frame, that follows, it are different, than it maybe a drastic motion change



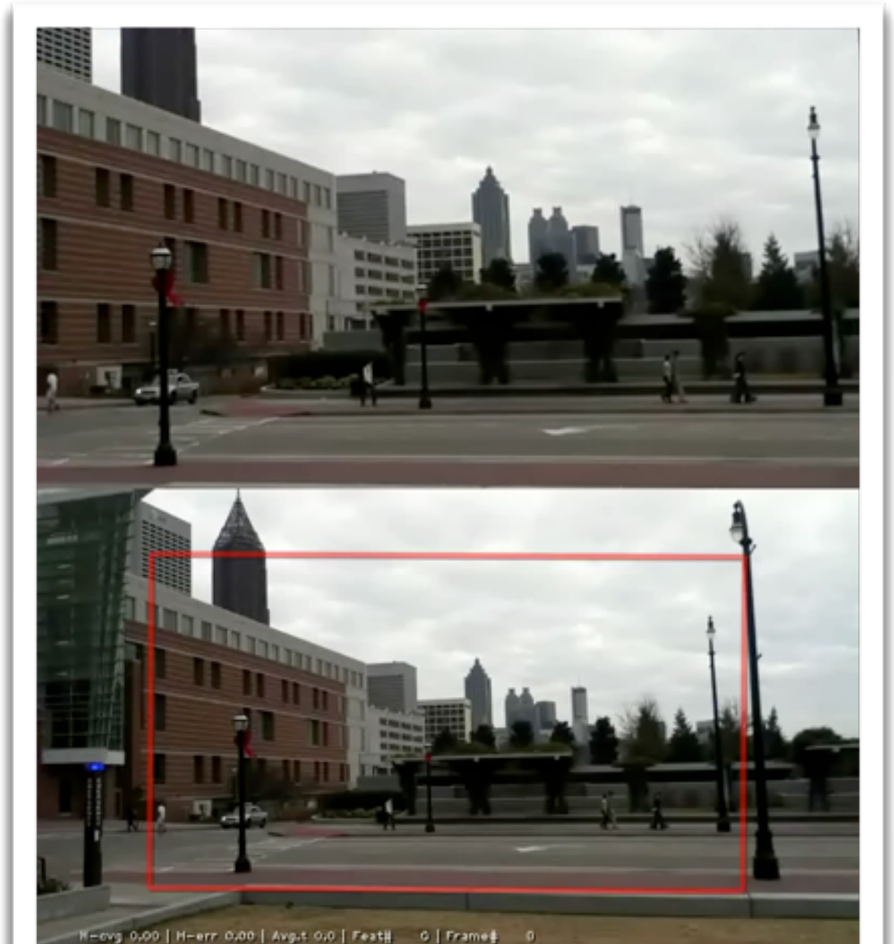
Feature Detection and Matching

* Same as in images
* Leverage the fact that features found in one frame may be visible in the next

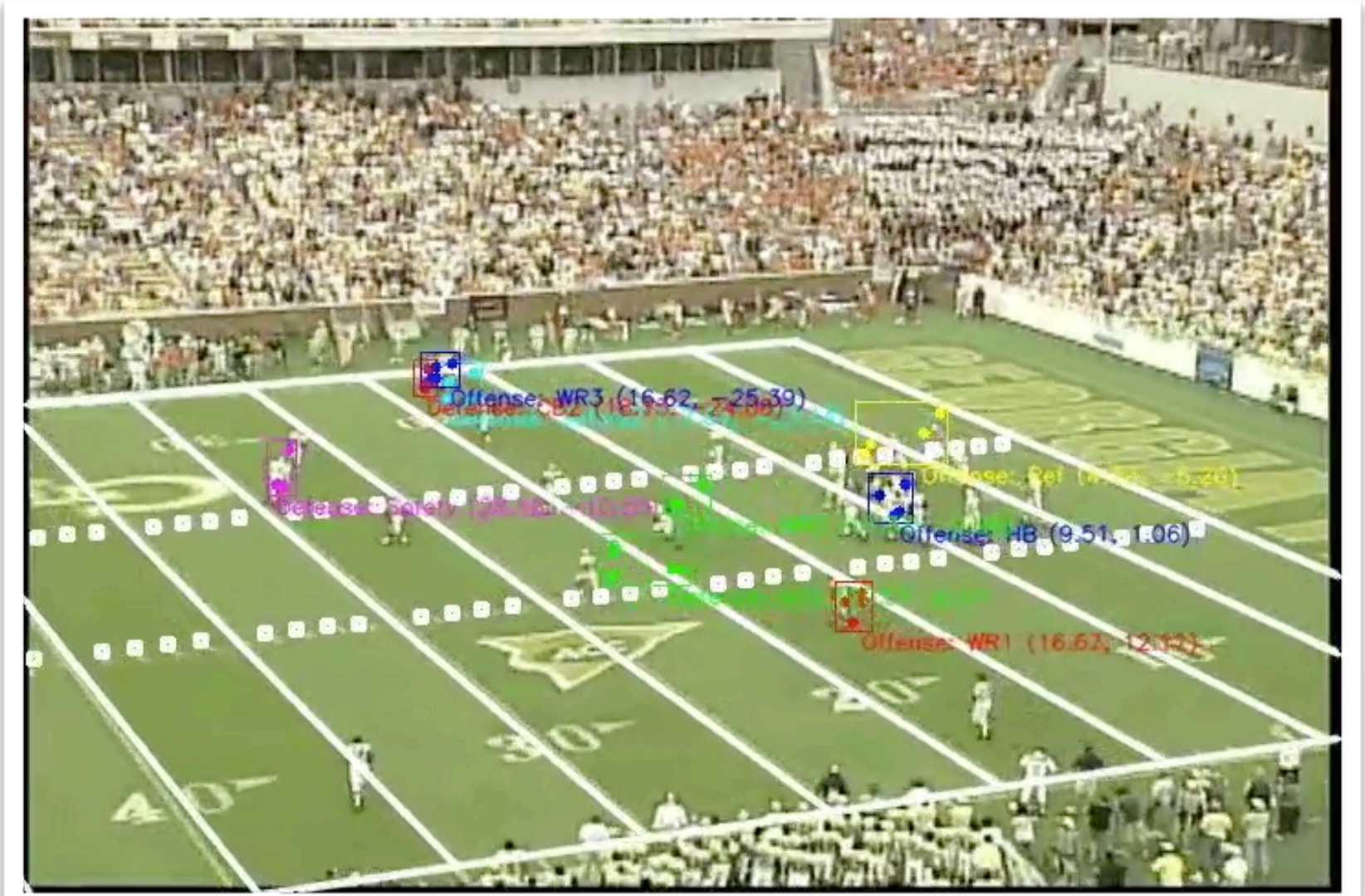


Feature Tracking

* Direct approaches to tracking
  * Find a feature, and match it to feature in the next frame)
* Motion-based approaches,
  * Compute the motion at pixel level between frames (OPTICAL flow)



Tracking, Registration in Video



(Hecht, Kazian, Mansour, St. John, Stallworth, Essa 07)

Registration and Blending in Video

(Kim, Oh, Essa 2009, 2011)

Summary

* Representational relationship between Images and Videos
* Persistence of vision in playing (and capturing) Videos
* Extension of filtering and processing from Images to Videos
* Methods used for tracking points in Videos

Next Class

* Video Textures and More . . .

Credits

* For more information, see
  * Richard Szeliski (2010) Computer Vision: Algorithms and Applications, Springer
* Some video retrieved from
  * http://commons.wikimedia.org/
  * From Professors Essa's Lab

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

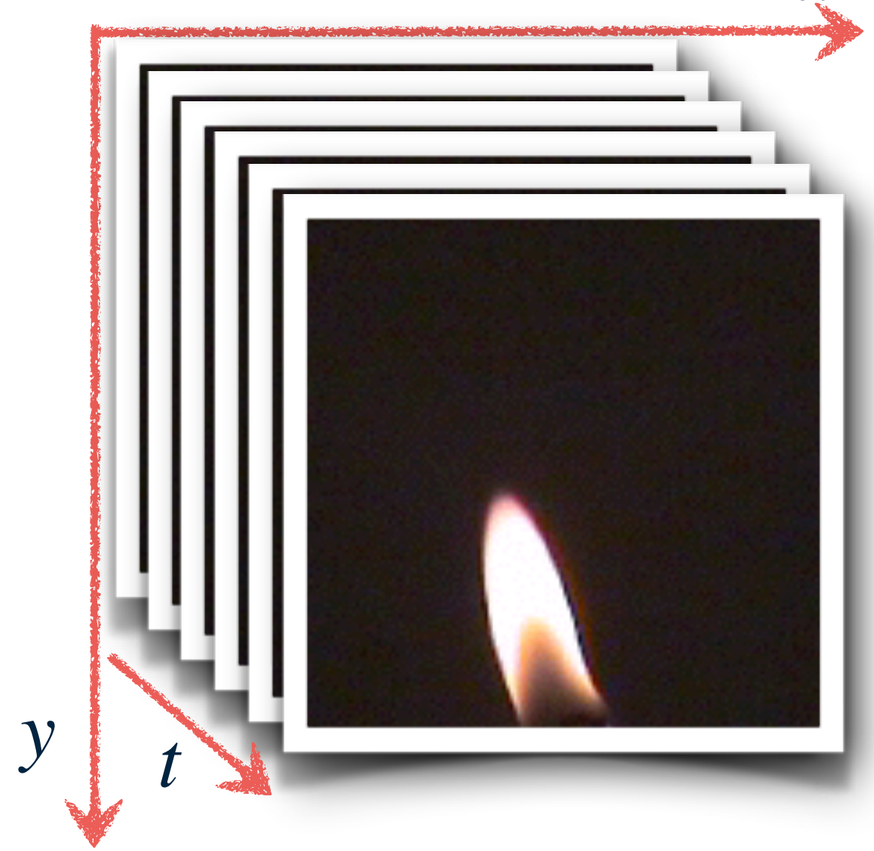
Georgia Tech College of Computing

Video Textures

* How can we find similar frames in a Video Volume to generate longer Videos?

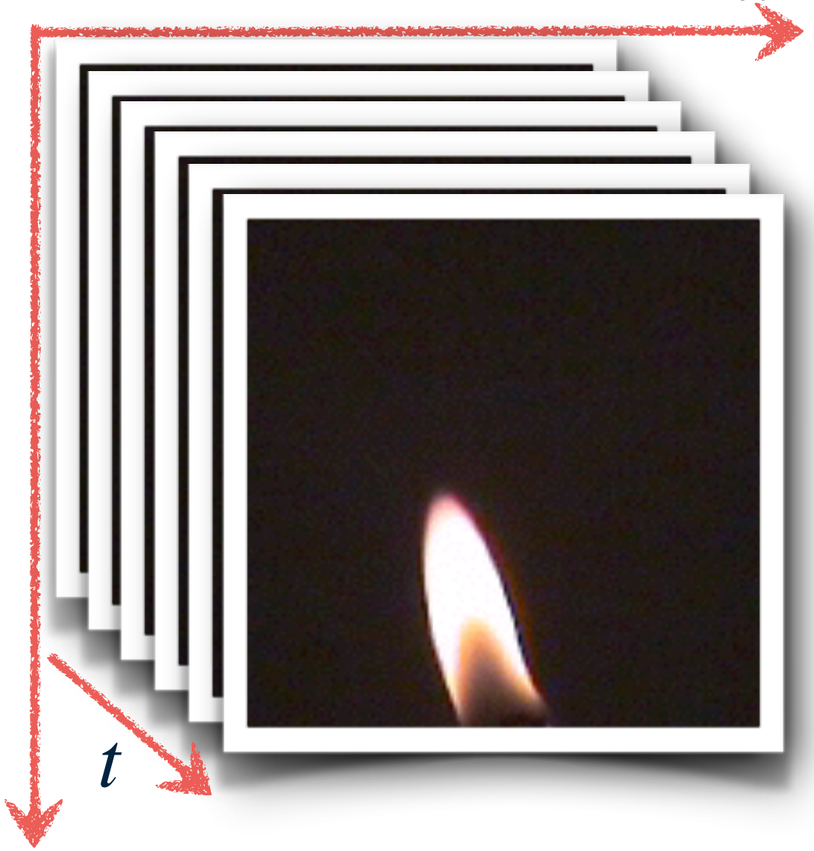
1. Concept of a Video Texture
2. Methods used to compute similarity between frames
3. Use of similar frames to find transitions to generate Video Textures
4. Fading, Cutting, Morphing for Video Textures
5. Some Applications of Video Textures

RECALL: Video is Images OVER time



* "Digital" Video:
  * numeric representation in two-dimensions (x and y), stacked in time, t
  * referred to as $I(x,y,t)$ in continuous function form, $I(i,j,t)$ in discrete

RECALL: Video is Images OVER time



* Video Resolution:
  * Expressed as representation of Width and Height of the image
  * Usually in aspect ratios of ***4x3***, ***16x9***, etc
* File formats: Include images, frame-rates, and codec/wrappers

Video Textures

Still Pictures

Schödl, Szeliski, Salesin, and Essa (2000)

Video Textures

Video

Video Textures

Looping Video

Video Textures

Video Textures

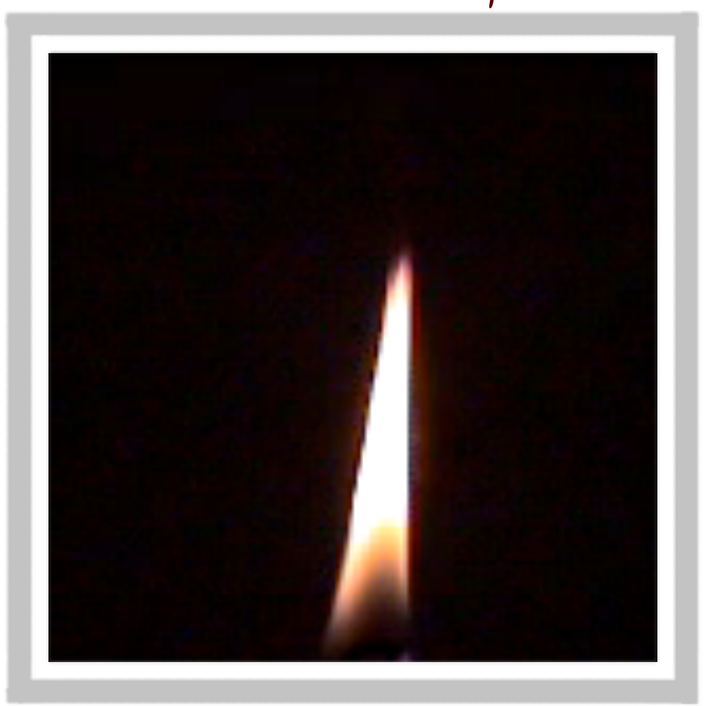
Video Clip to Video Textures

Video Clip

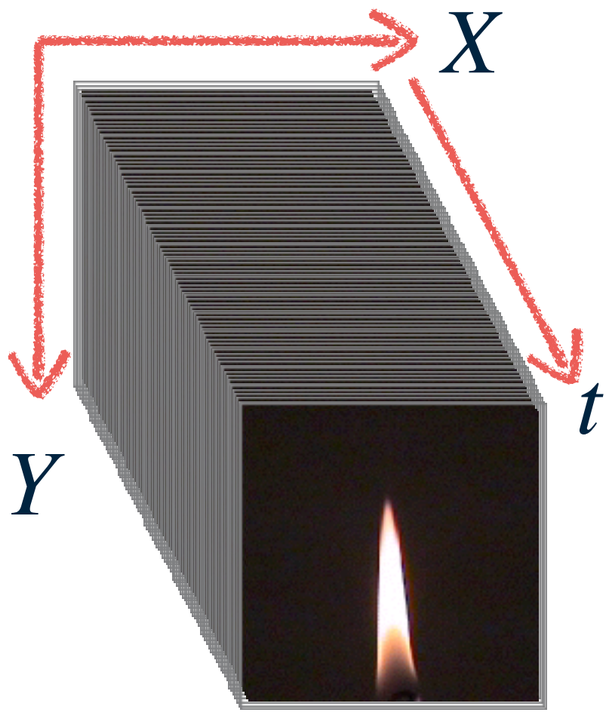
Video Texture

Video Clip to Video Textures

Video Clip



90 frames



90 frames:

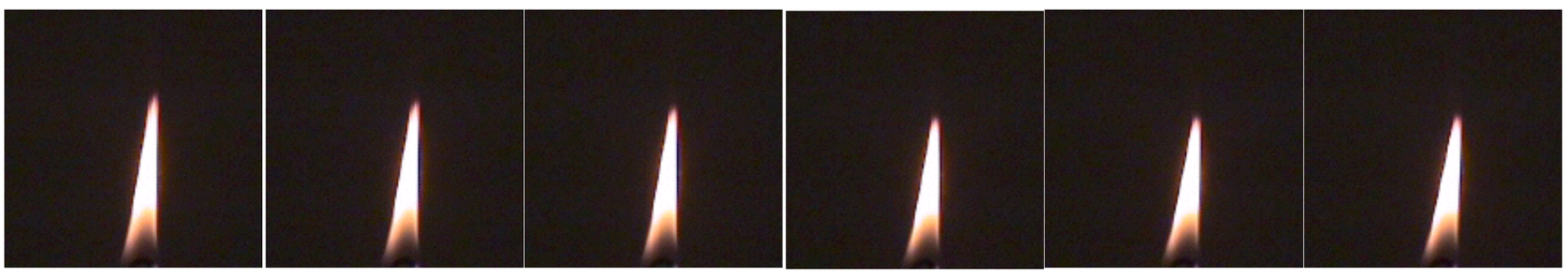
$f_1, f_2, f_3, \ldots, f_{90}$

f_1

Compute how similar f_1 is to all frames: $f_1, f_2, f_3, \ldots, f_{90}$

Do this for all $f_1, f_2, f_3, \ldots, f_{90}$

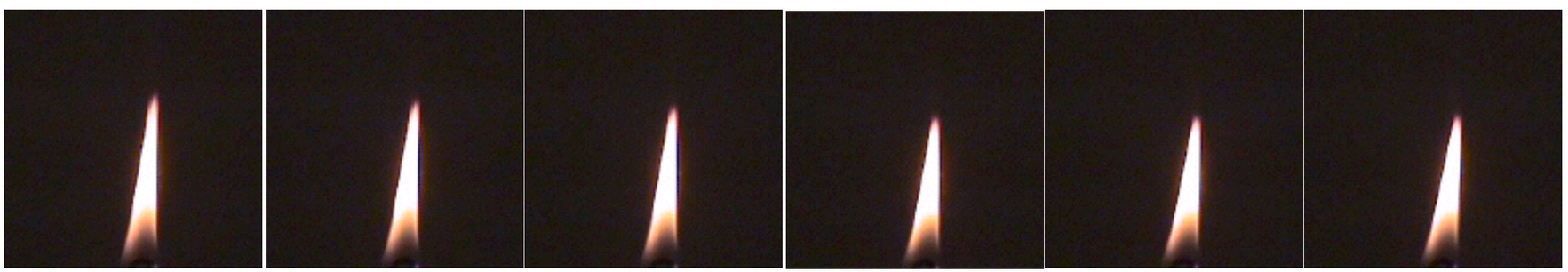
Similarity Metric (1)



f_1 f_2 f_3 f_4 f_5 f_6

* Compute the Euclidean Distance between two Frames

Similarity Metric (1)



f_1 f_2 f_3 f_4 f_5 f_6

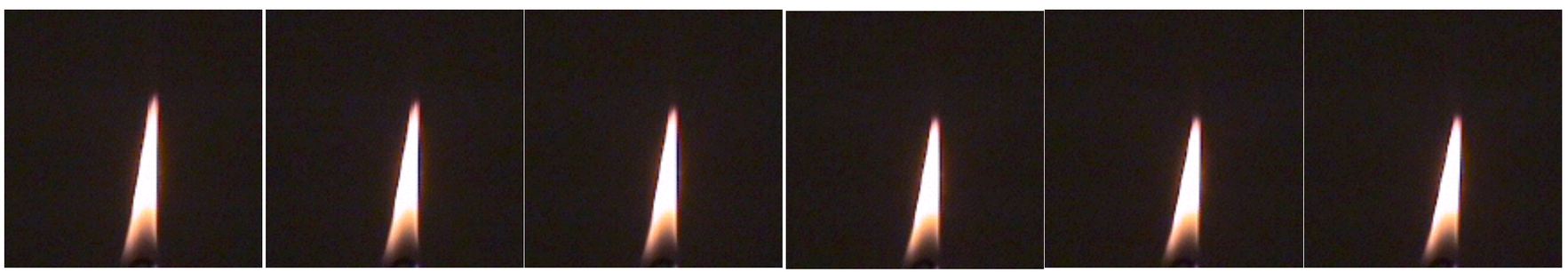
* Consider two frames, $p = \{p_1, p_2, \ldots p_n\}$ and $q = \{q_1, q_2, \ldots q_n\}$

$$d_2(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

$$d_2(p,q) = \sqrt{\sum_{i=1}^{N} (p_i - q_i)^2}$$

This distance metric is referred to as L2 norm

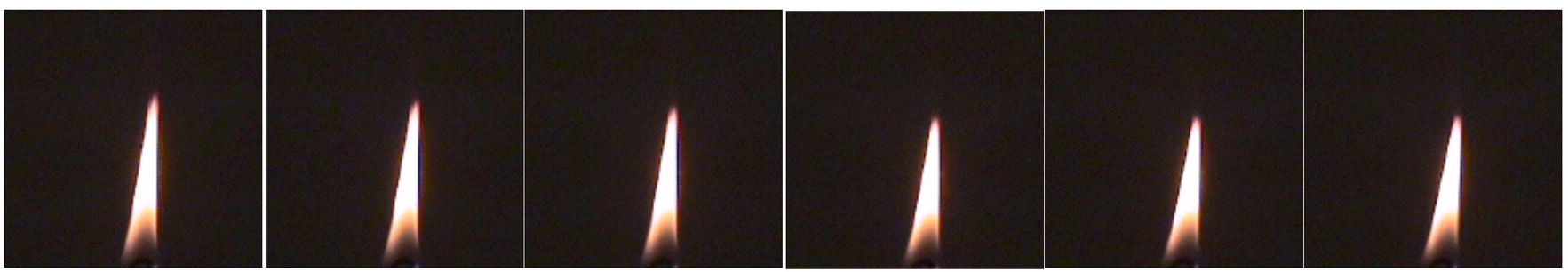
Similarity Metric (2)



f1 f2 f3 f4 f5 f6

* Compute the manhattan Distance between two Frames

Similarity Metric (2)



f_1 f_2 f_3 f_4 f_5 f_6

* Consider two frames, $p = \{p_1, p_2, \ldots p_n\}$ and $q = \{q_1, q_2, \ldots q_n\}$

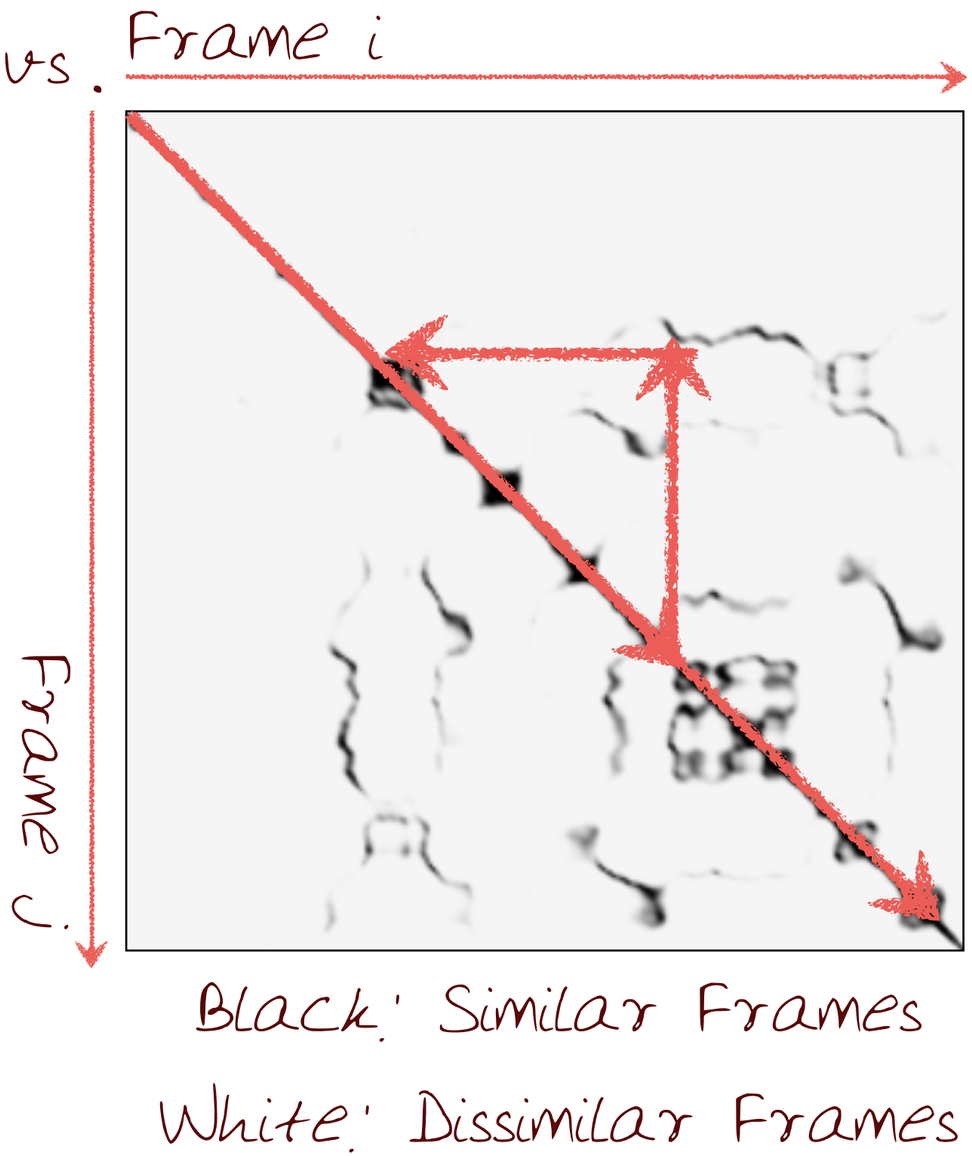
$$d_1(p, q) = (p_1 - q_1) + (p_2 - q_2) + \cdots + (p_n - q_n)$$

$$d_1(p, q) = \sum_{i=1}^{N} |p_i - q_i|$$

This distance metric is referred to as L1 norm. (|..| implies Absolute)

Finding Similar Frames

vs.



Black: Similar Frames

White: Dissimilar Frames

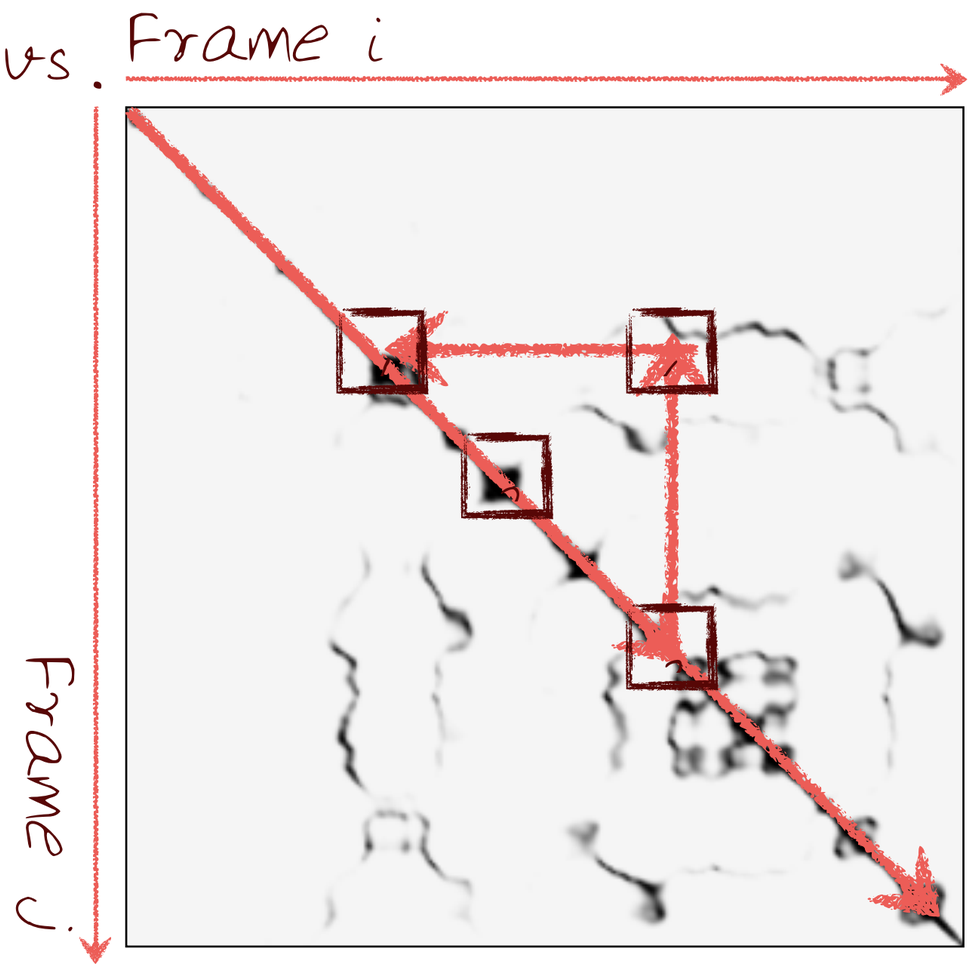
* Compute Euclidean Distance d_{ij} between all N Frames

* Similar frames are the ones that would be best to jump to

Infinitely Long Video Texture

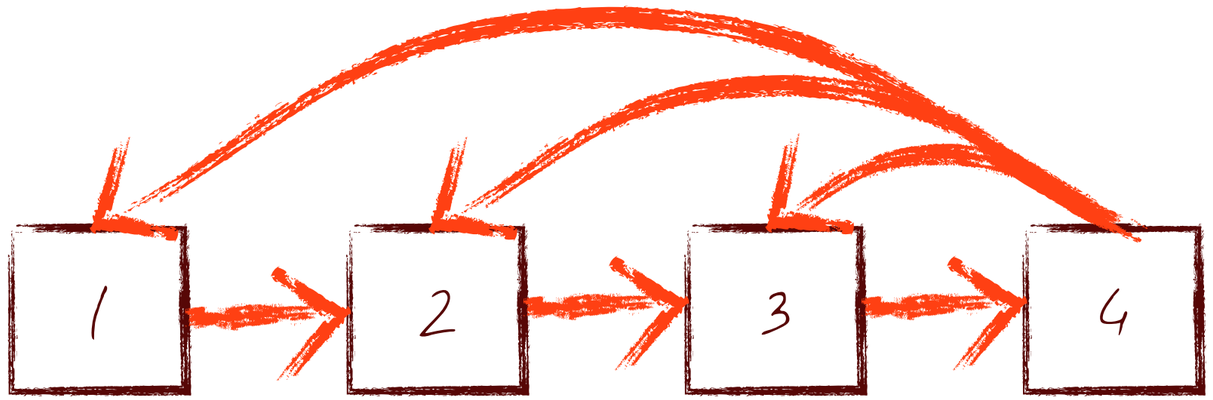
Finding Similar Frames

vs.



Black: Similar Frames

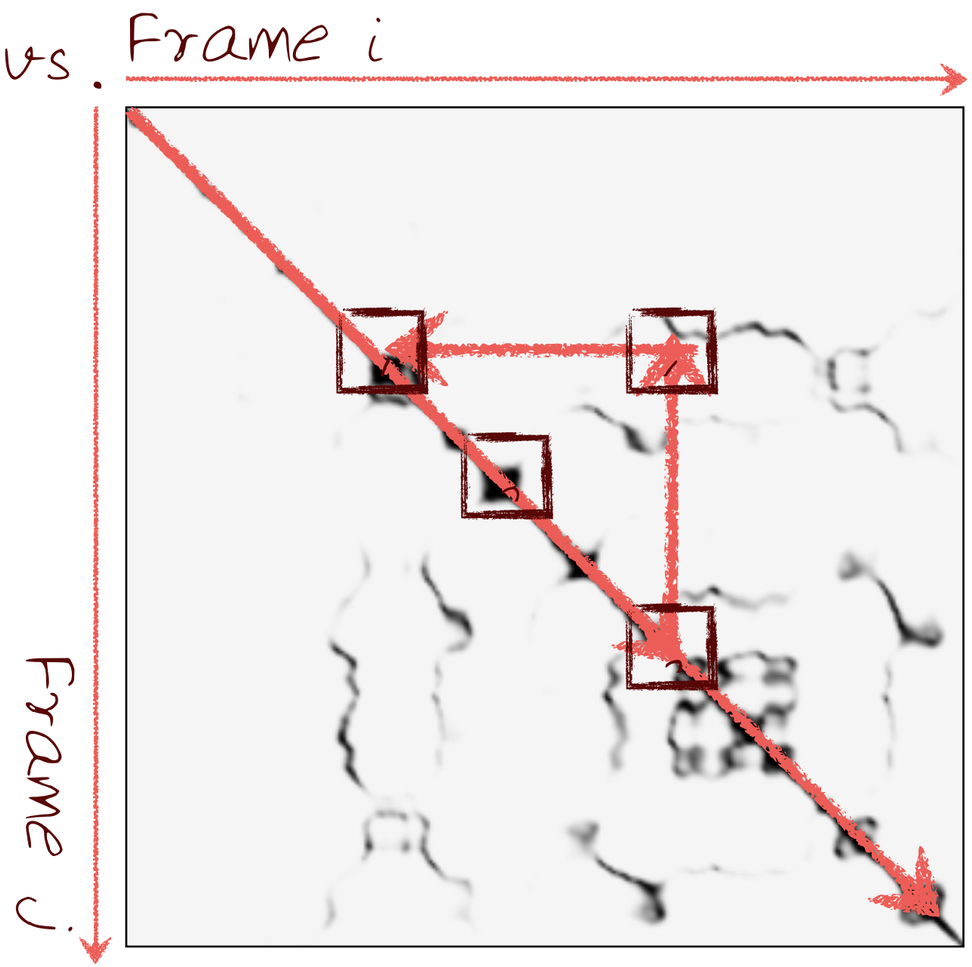
White: Dissimilar Frames



Markov Chain

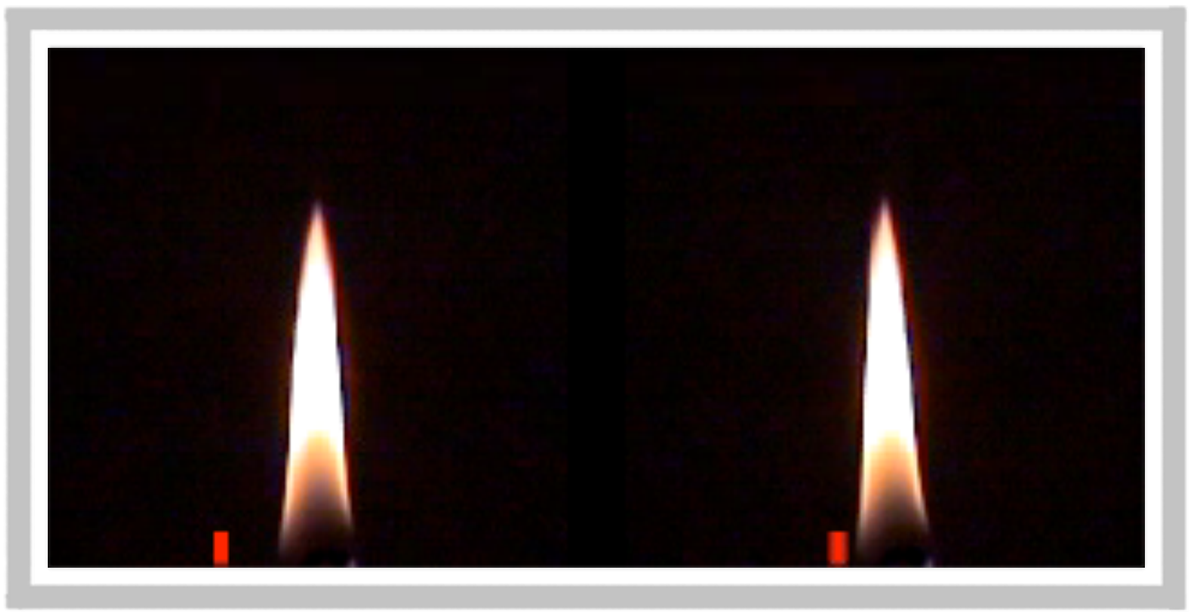
Finding Similar Frames

vs.

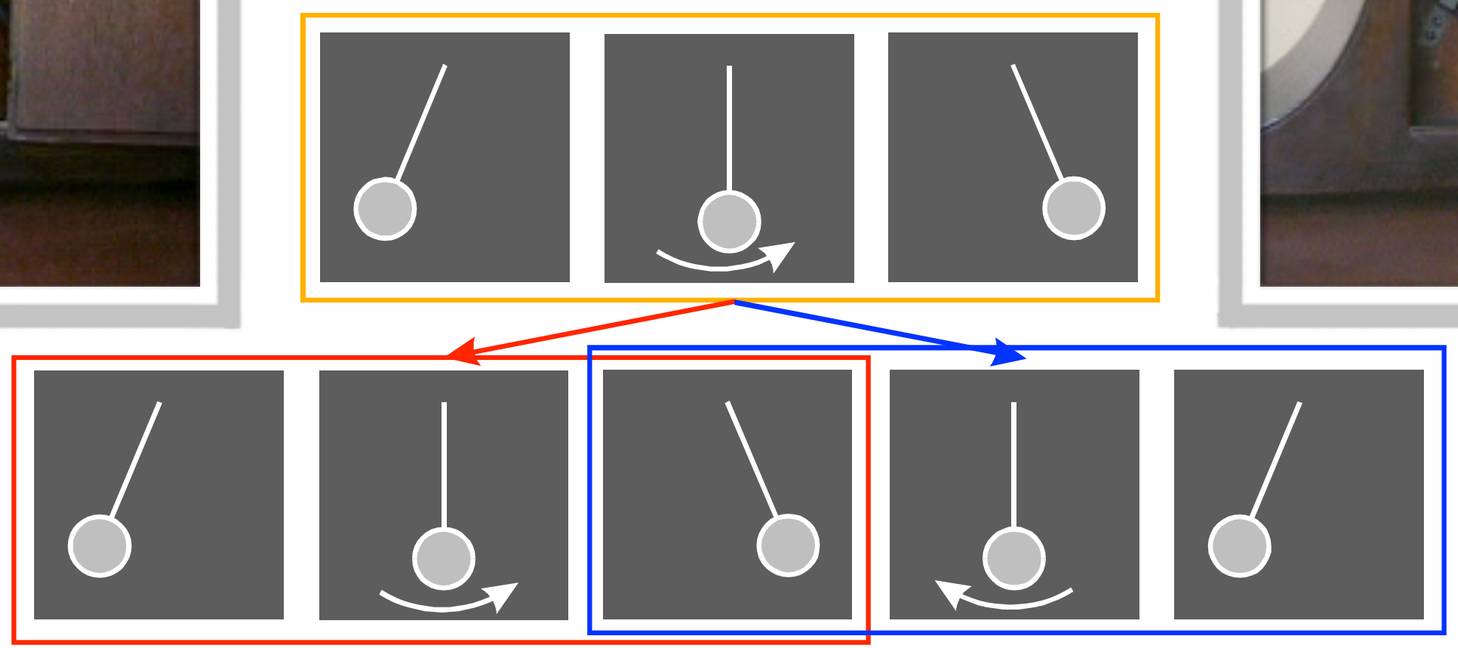


Black: Similar Frames

White: Dissimilar Frames



Preserving Dynamics with Transitions



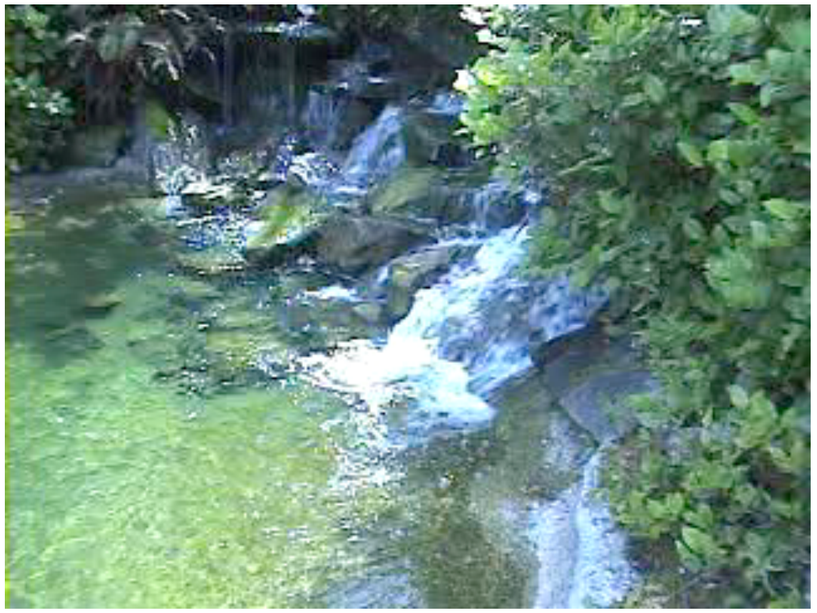
Fading and Blending in Video

Cut

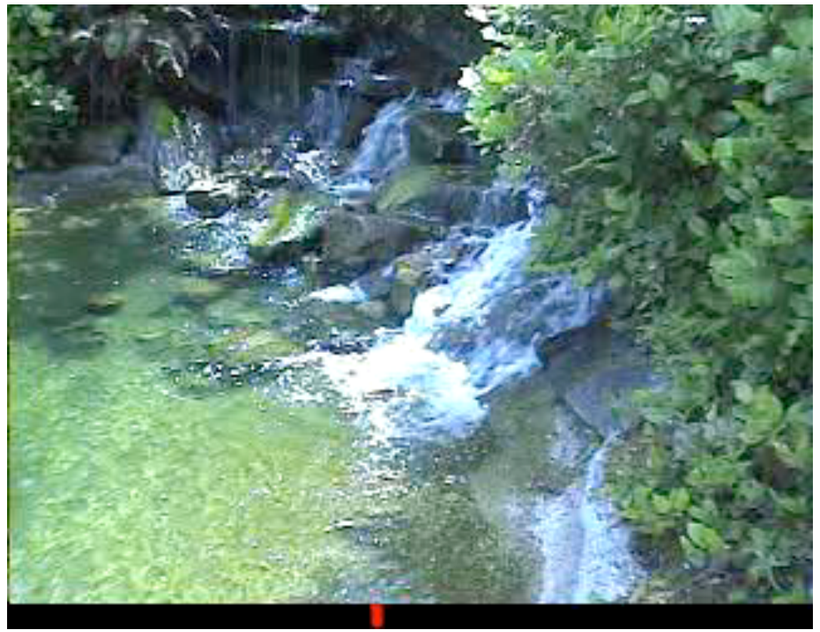
Fade

Morph/Feather

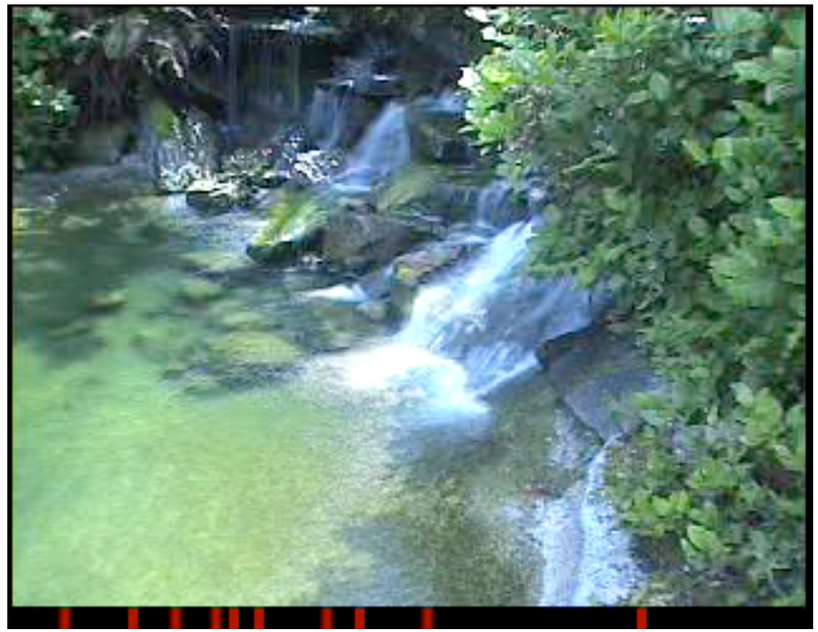
Fading and Blending in Video



Original

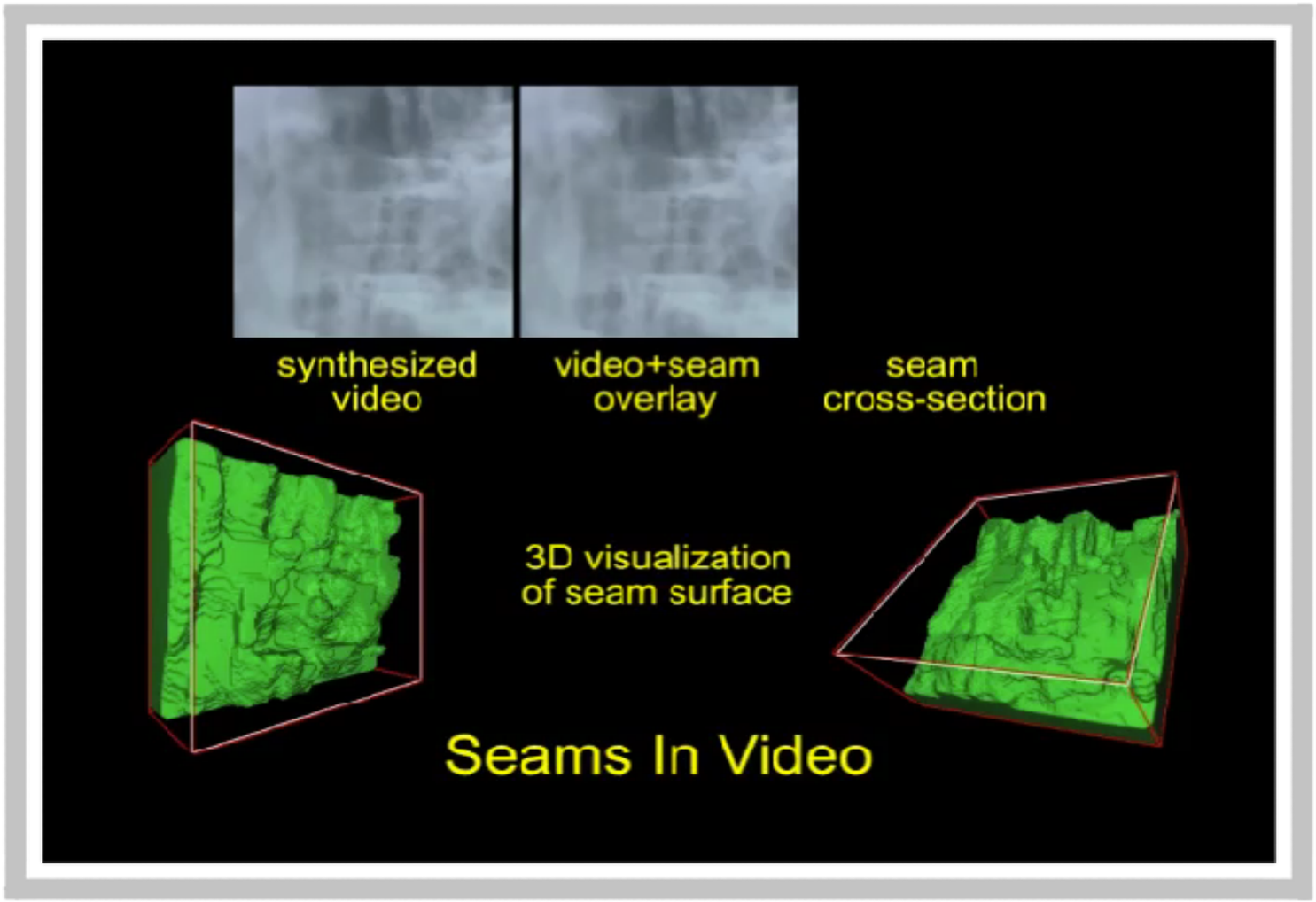


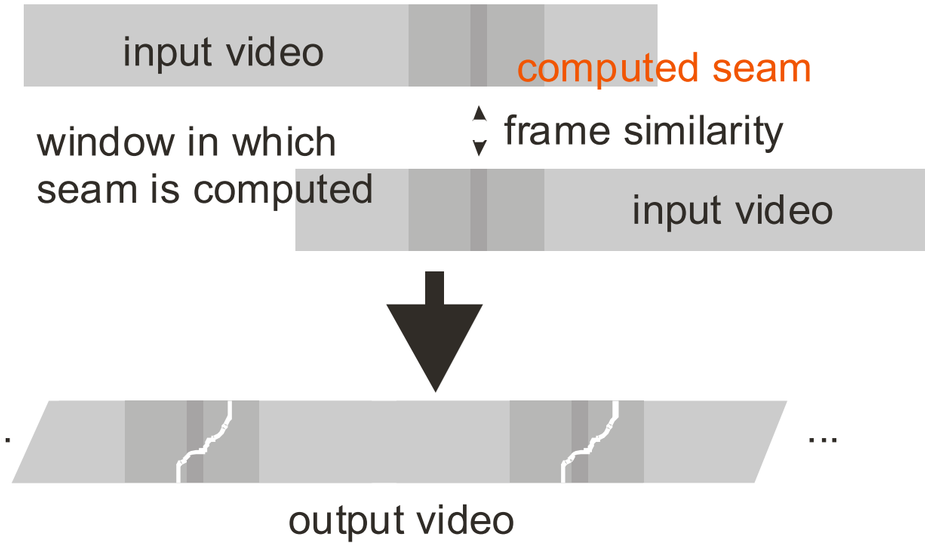
Single Fade



Multiple Fade/Blends

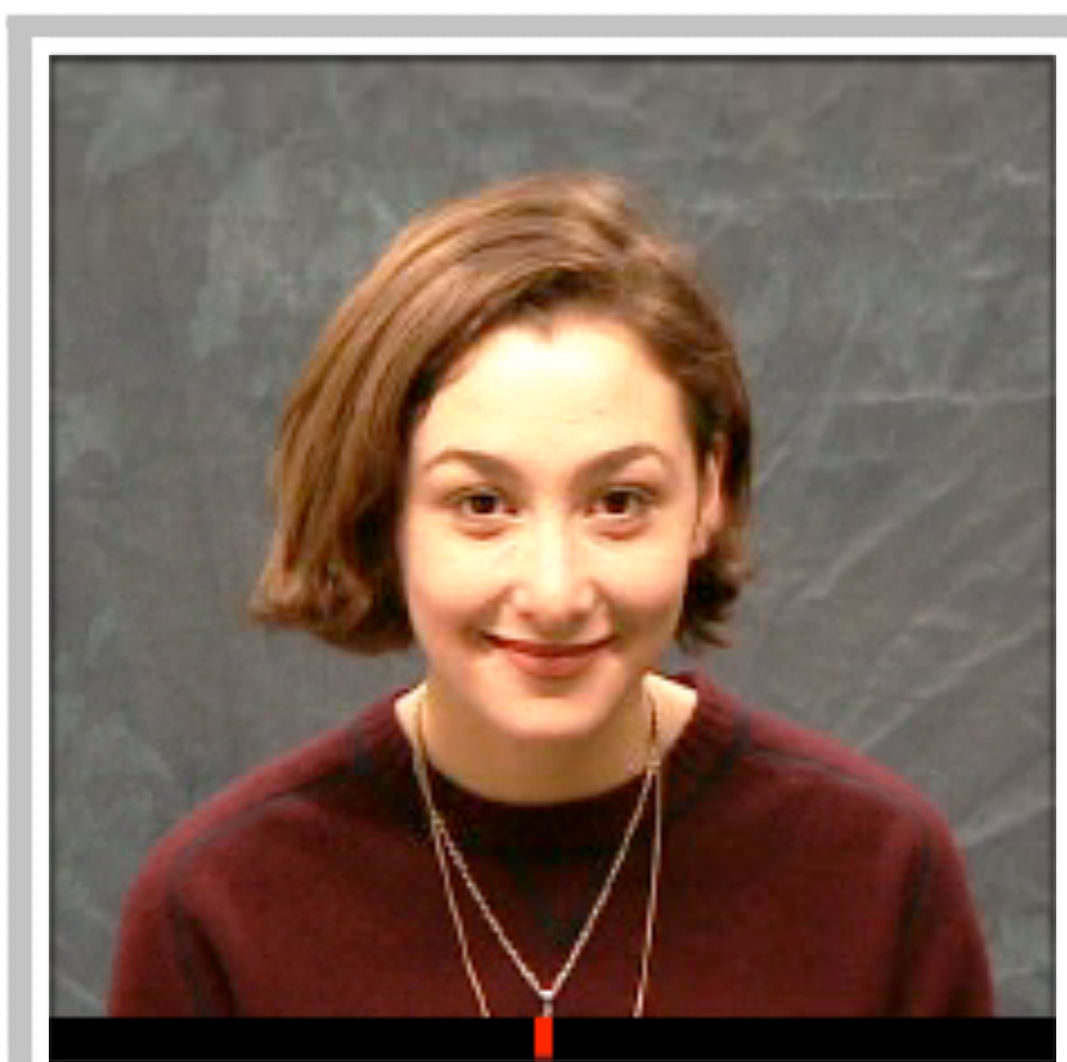
Not Just Fade/Blend, but Cut

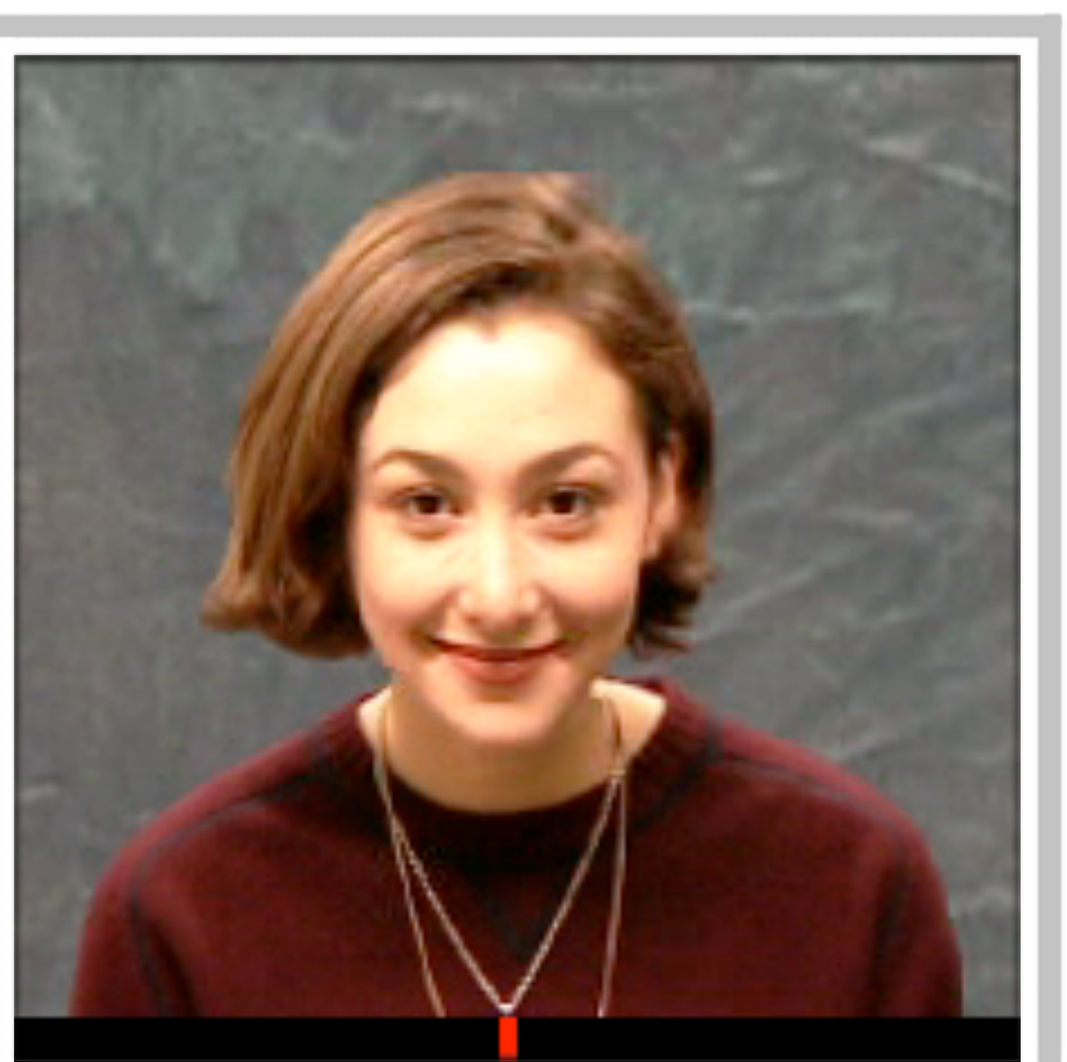




(Kwatra, Schödl, Essa 2003)

Video Portraits

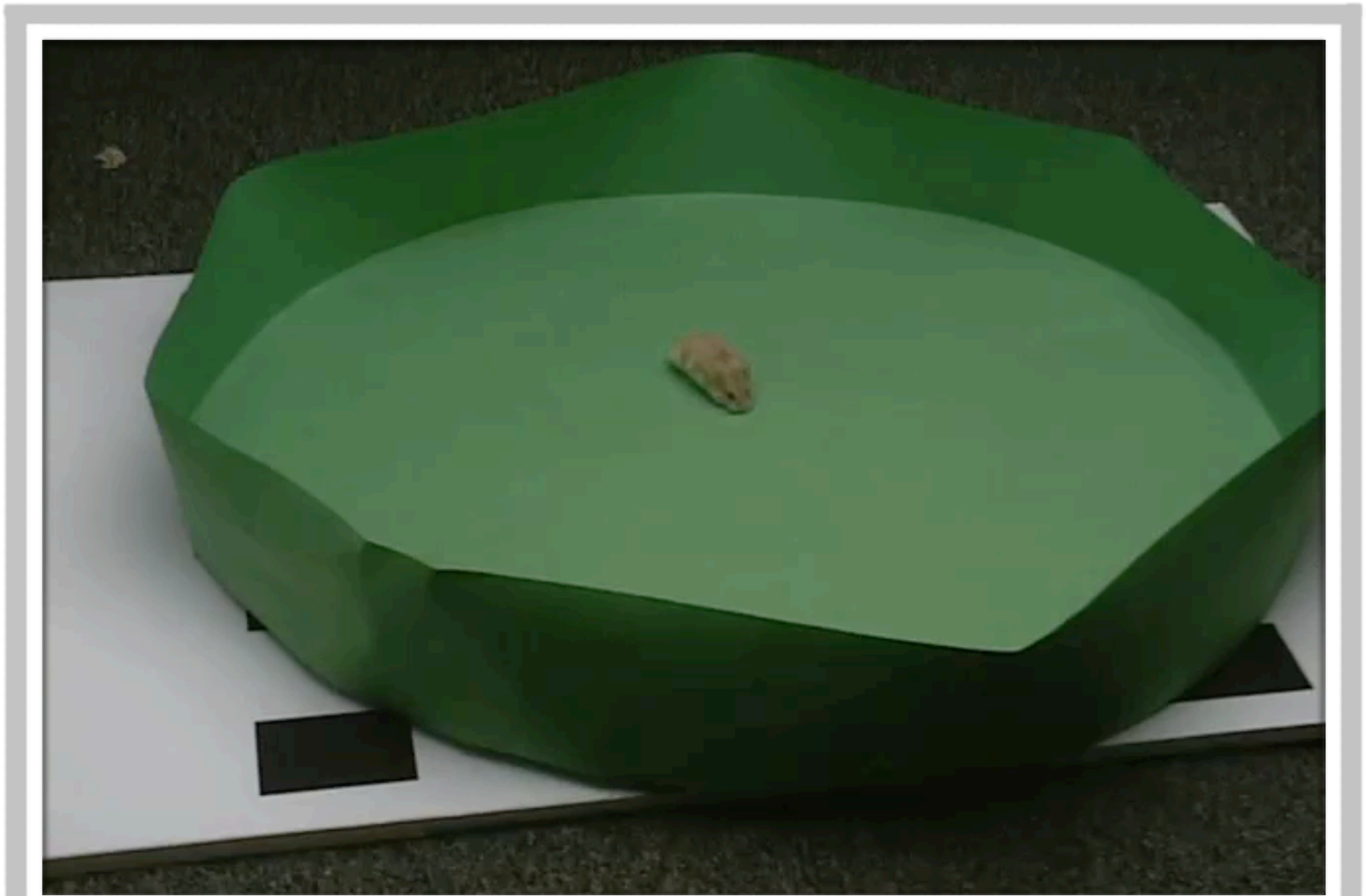




Video Sprites

Schödl and Essa (2002)

Video Sprites



Schödl and Essa (2002)

Cliplets Cinemagrams

Selecting and playing (animating) only parts of a Video

Henrik Sozzi,
http://www.youtube.com/watch?v=gkYL9weEDWE
using Microsoft Research Cliplets

Tom Freestone,
http://www.youtube.com/watch?v=u_scohnE9Wg

Summary

* Introduced the concept of a Video Texture
* Methods used to compute similarity between frames
* Using similar frames to generate Video Textures by finding similar points to transition to
* Discussed blending, fading, and cuts to generate Video Textures
* Extensions of Video Textures

Further Information

* Schödl, Szeliski, Salesin, and Essa (2000), "Video textures," in SIGGRAPH 2000
* Kwatra, Schödl, Essa, Turk, Bobick (2003), "Graphcut textures: image and video synthesis using graph cuts" in SIGGRAPH 2003
* Schödl and Essa (2002), "Controlled animation of video sprites" in ACM SIGGRAPH Symposium on Computer animation
* Agarwala, Zheng, Pal, Agrawala, Cohen, Curless, Salesin, and Szeliski (2005) "Panoramic video textures." SIGGRAPH 2005
* Bai, Agarwala, Agrawala, Ramamoorthi (2012), "Selectively De-animating Video," SIGGRAPH 2012

Credits

* For more information, see
  * Richard Szeliski (2010) Computer Vision: Algorithms and Applications, Springer
* Some videos retrieved from
  * From Professors Essa's Lab
  * Microsoft Research Cliplets

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Video Stabilization

* Stabilization (removing excessive shake/motion) in videos

Georgia Tech College of Computing

Lesson Objectives

1. Video stabilization
2. Estimating camera motion
3. Smoothing camera paths
4. Rendering stabilized videos
5. Dealing with Rolling Shutter artifacts

Casual Video (Original)

Casual Video Stabilized

Stabilized Videos (Side by Side)

Original

Stabilized

Video Stabilization

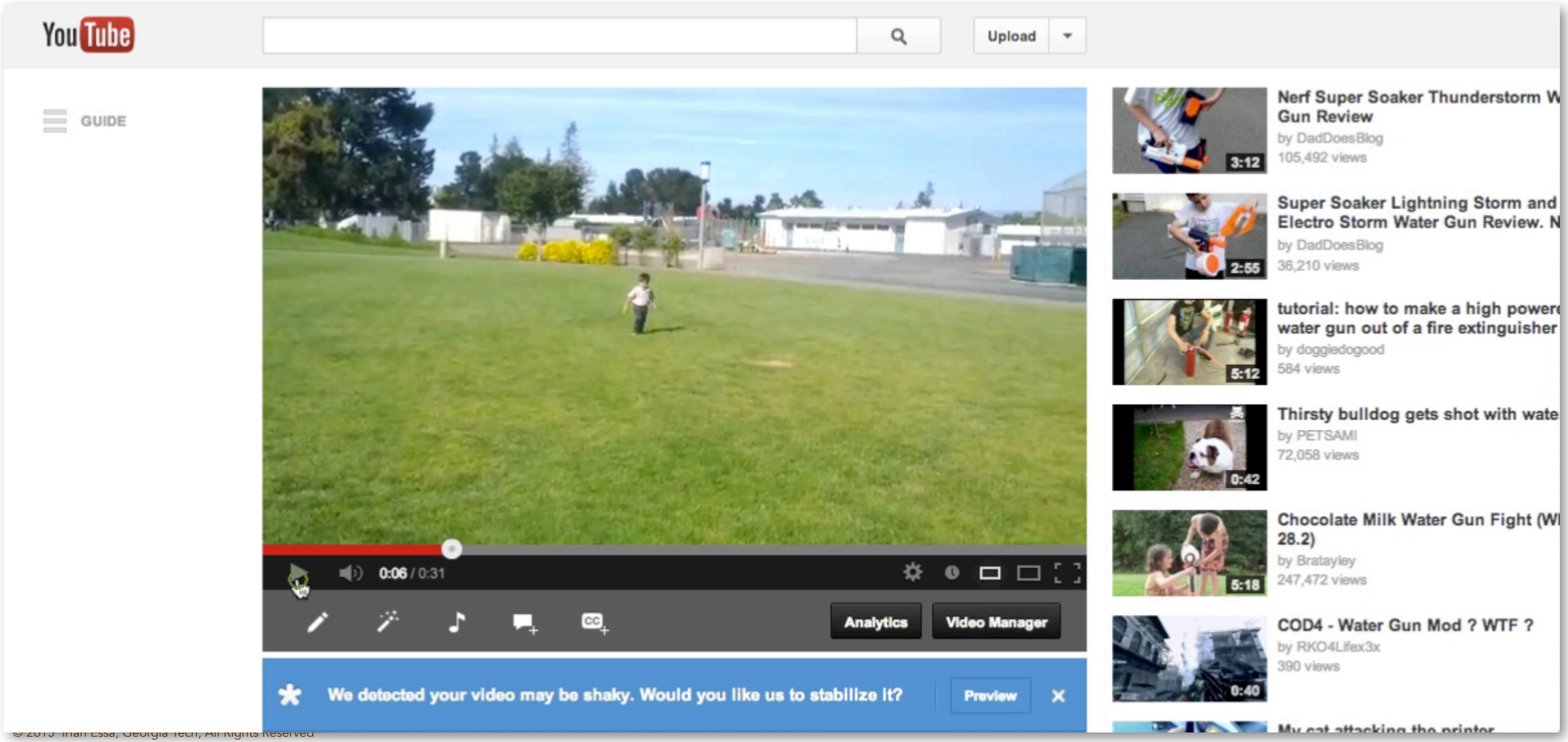
Georgia Tech

Research at Google

* Grundmann, Kwatra, and Essa (2011) IEEE CVPR 2011
* Grundmann, Kwatra, Castro, and Essa (2012) in IEEE ICCP 2012
* Google Team: Doshi, Bridgwater, Kwan, deLespinasse Steger, Glickstein, Toff

YouTube Enhancement Suite

YouTube Shake Detection and Removal



Video Stabilization Types

Steadicam (1975) by inventor and cameraman Garrett Brown

http://en.wikipedia.org/wiki/Steadicam

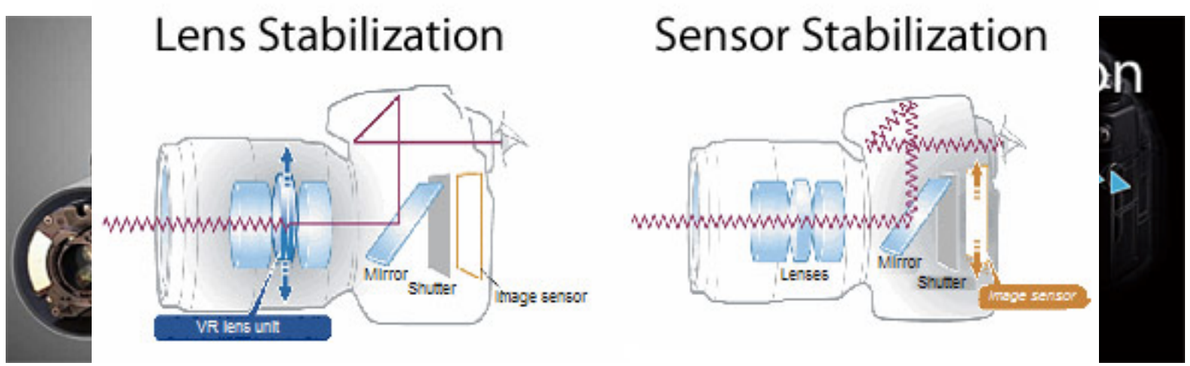
'Information Overdrive!' Tips For Professional Video Recording With A Steadicam - Crews Control, Inc

Recall: Evolution of the Camera

Video Stabilization Types

Optical / In-camera Stabilization

- Floating lens (electromagnets)
- Sensor shift
- Accelerometer + Gyro
- High-frequency perturbations (small buffer)



Post-process Stabilization

- Removes low-frequency perturbations (large buffer)
- Distributed backend processing (cloud computing)
- Can be applied to any camera, any footage

Post-process Video Stabilization

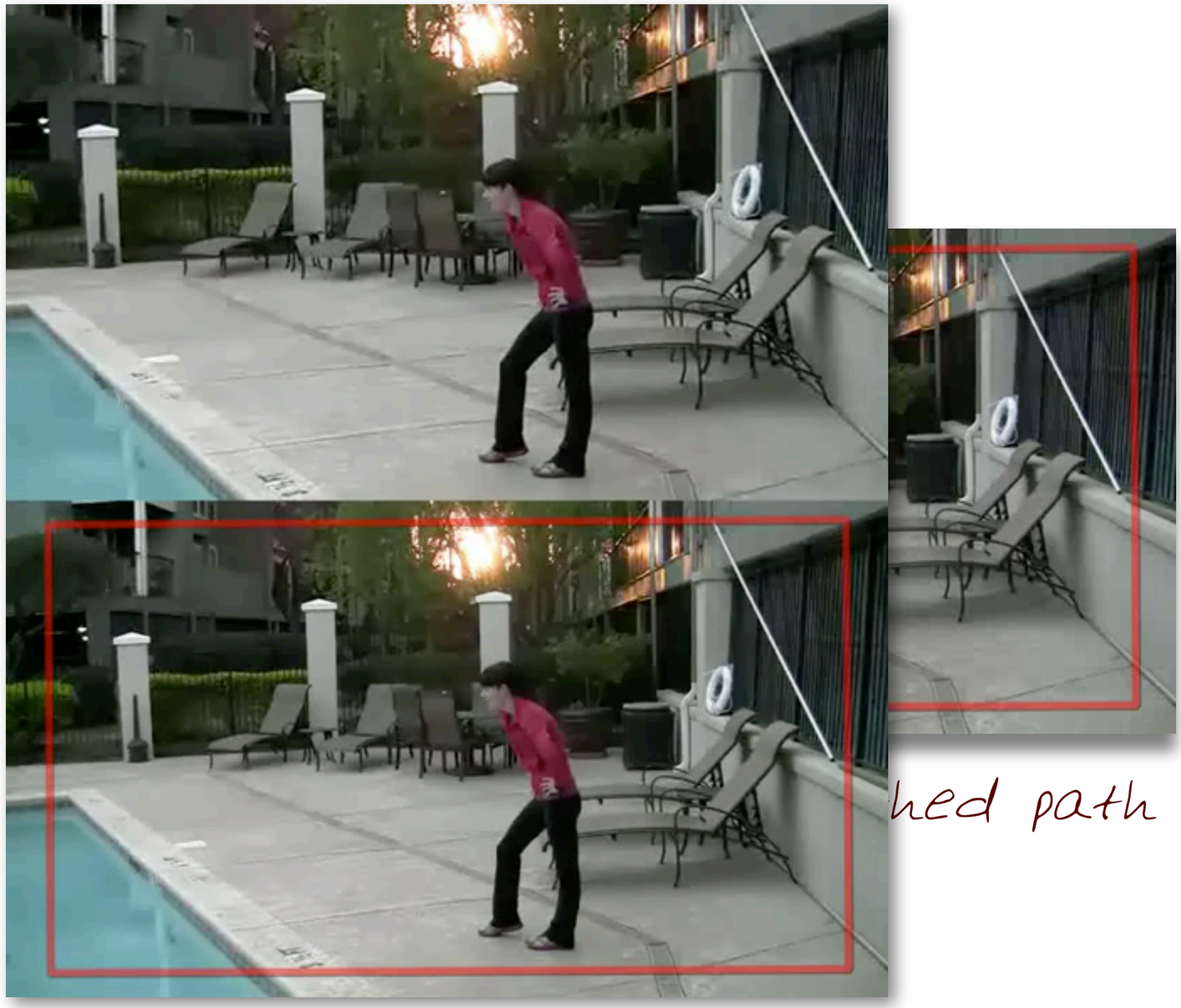
Main Steps

(1) Estimate Camera Motion → (2) Stabilize Camera Path → (3) Crop and Re-synthesize

Original video (shaky)

Stable, Virtual Camera

* Challenges:
  * Can deviate too much from original camera
  * Undefined content (black borders)

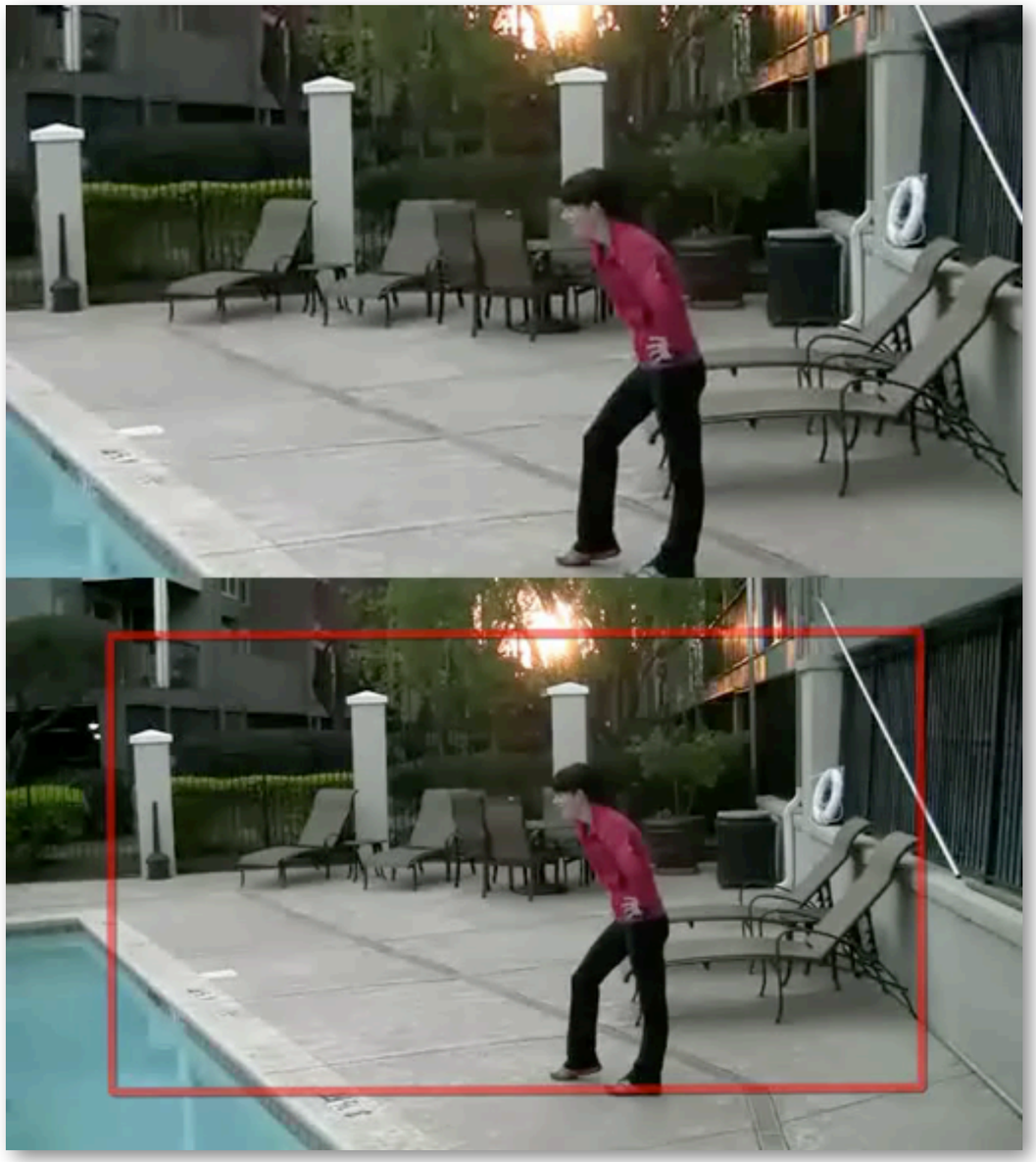


hed path

Stabilization By Cropping

* Solution:
  * Constrain crop to stay within frame bounds

* Guarantee:
  * Never undefined content, avoids borders and in-painting



Post-process Video Stabilization

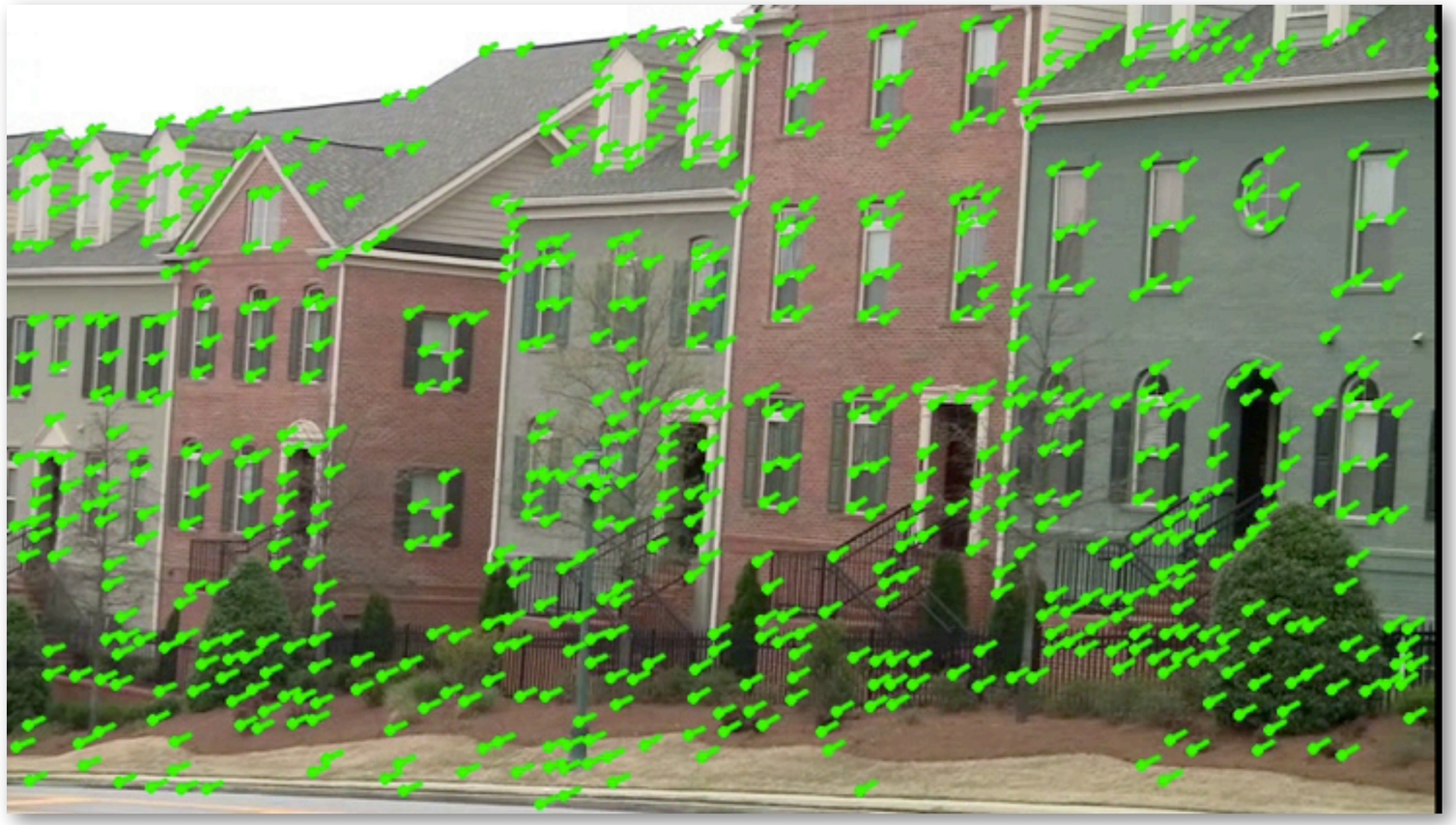
main Steps

Camera Path Estimation

(1) Estimate Camera Motion

(2) Stabilize Camera Path

(3) Crop and Re-synthesize



- Find image corners (high gradient in x & y)
- Track w.r.t. the previous frame

Background Motion

(1) Estimate Camera Motion

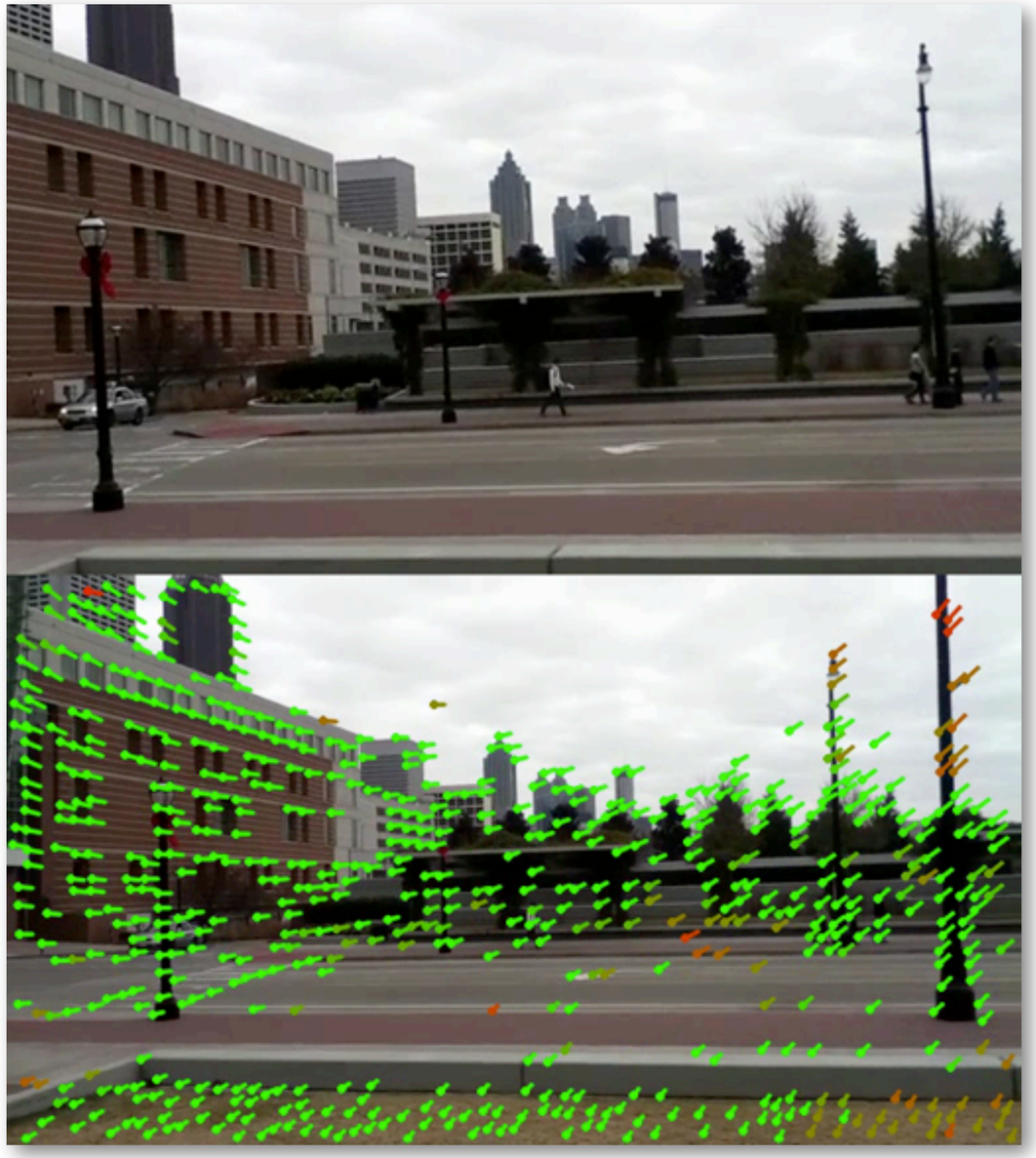
(2) Stabilize Camera Path

(3) Crop and Re-synthesize

* Only estimate camera motion of background

* Model contribution to background by weighting features

Background Foreground



Motion Models

(1) Estimate Camera Motion

(2) Stabilize Camera Path

(3) Crop and Re-synthesize

* Goal: Describe camera motion with fewer degree of freedoms (DOF)

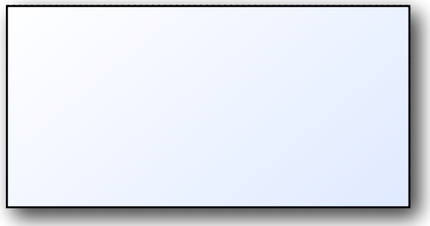
Motion Models

1. Translation

* Translation in x and y
* 2 DOF
* Still very shaky

Motion Models

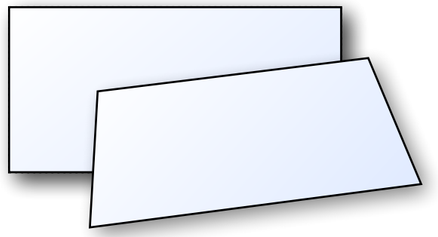
2. Similarity



* Translation in x and y
* Uniform scale and rotation
* 4 DOF
* Not shaky, but wobbly

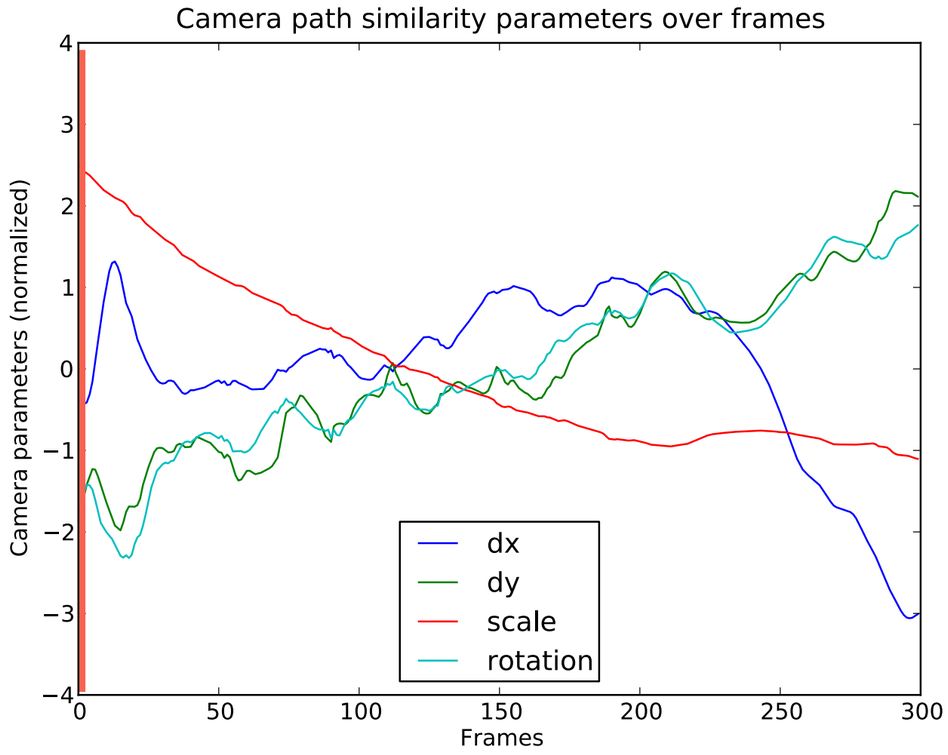
Motion Models

3. Homography



* Translation in x and y, scale and rotation
* Skew and perspective
* 8 DOF
* Stable

Similarity Model Over Time



* Four degrees of freedom (DOF):
  * Translation dx
  * Translation dy
  * Scale
  * Rotation

Smoothing Camera Paths

Camera Path Estimation

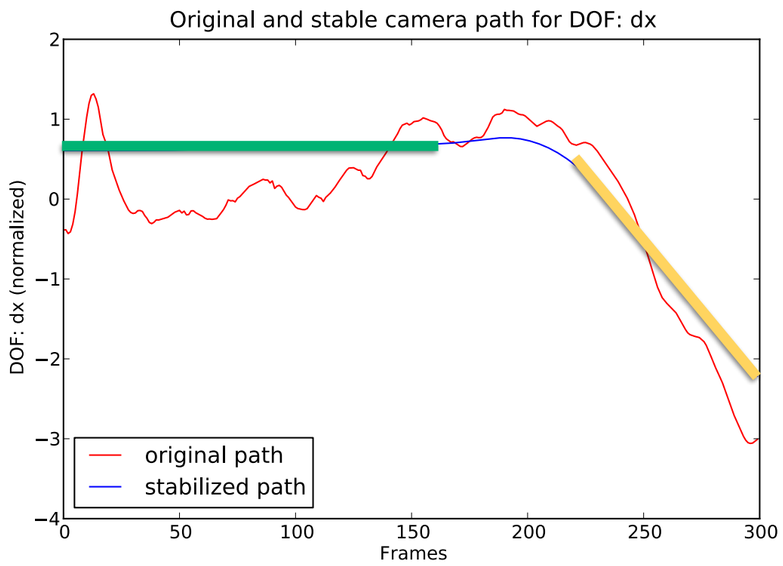
(1) Estimate Camera Motion

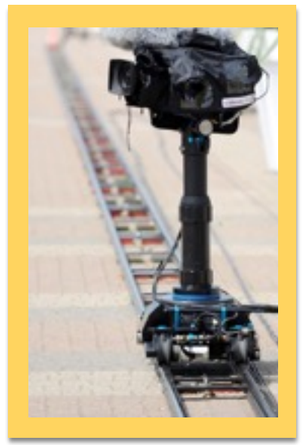
(2) Stabilize Camera Path

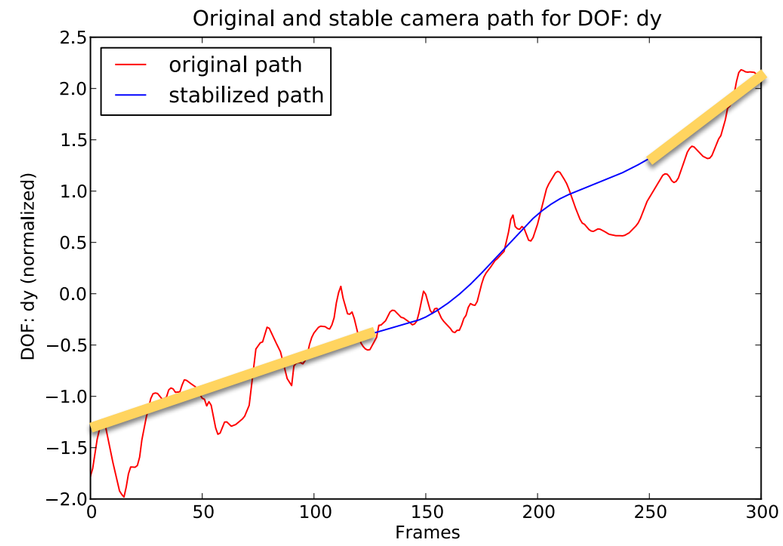
(3) Crop and Re-synthesize

Path Smoothing

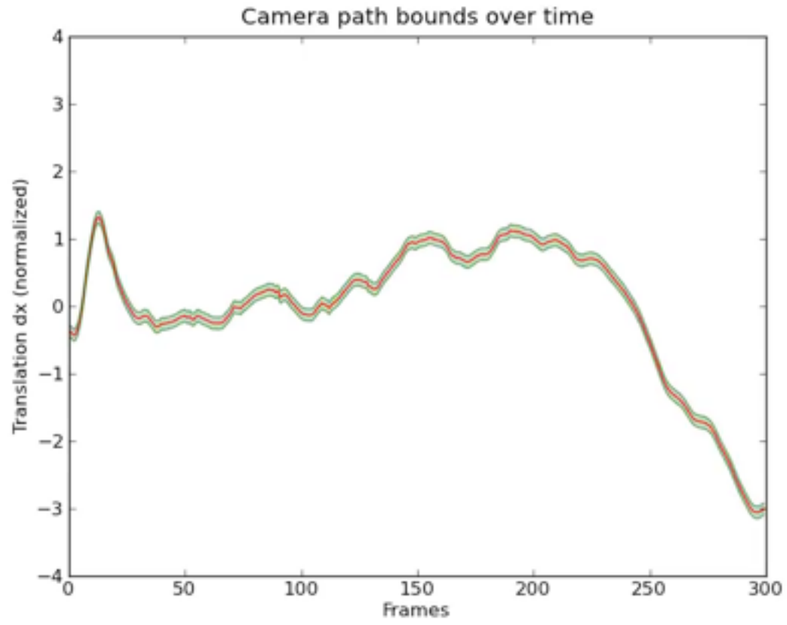
- Goal: Approximate original path with stable one
- Cinematography inspired: Properties of a stable path?
  - Tripod → Constant segment
  - Dolly or pan → Linear segment
  - Ease in and out transitions → Parabolic segment
- Solution: Find constrained partition





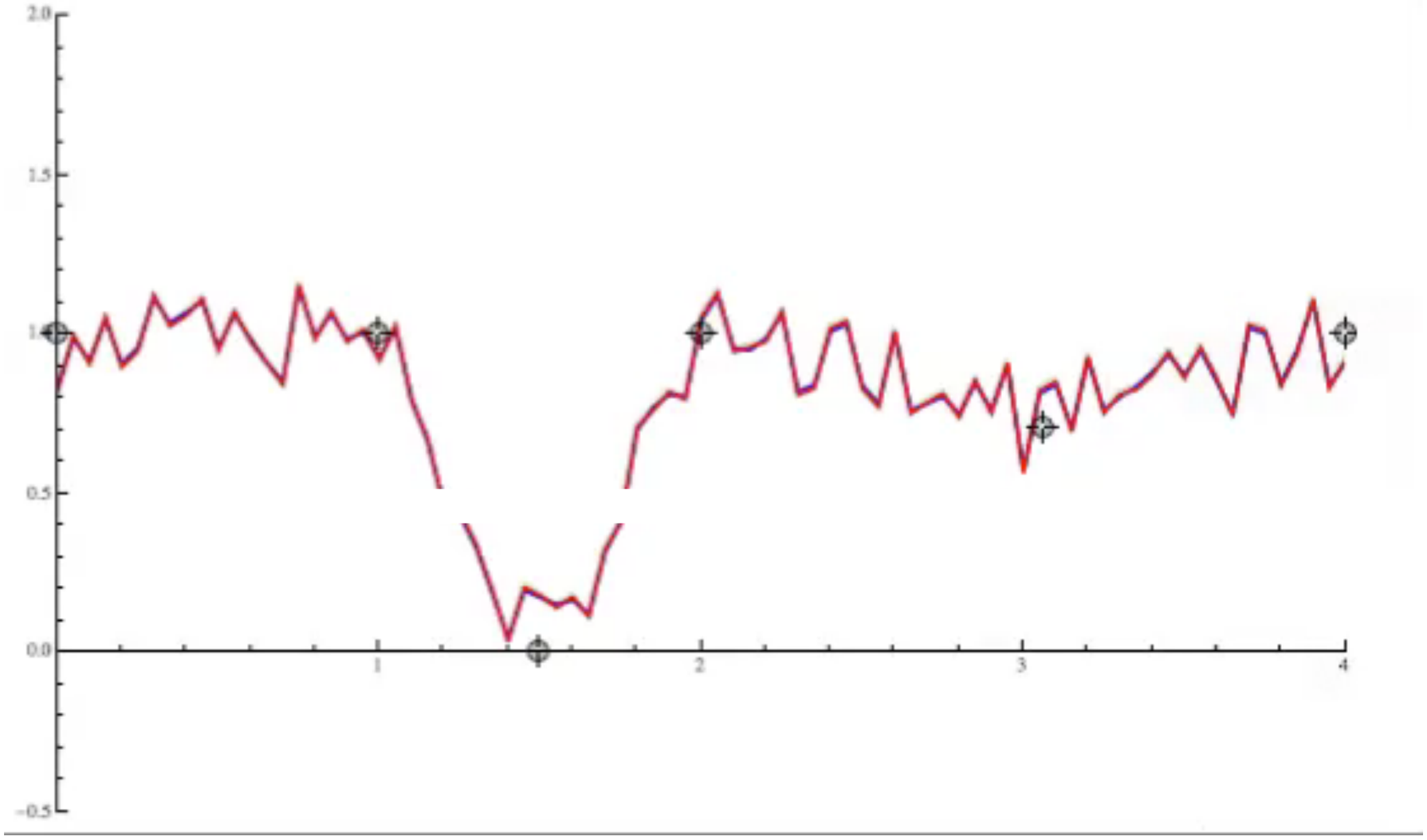


Path Smoothing



- Important constraint: Crop window within frame
- Crop window size = Envelope around original camera path
- Within the envelope: Find partition of constant, linear and parabolic segments

Path Smoothing Demo



Increase bounds Robust to perturbations Constant paths Linear paths Parabolic paths YouTube paths

Smoothing Camera Paths

Re-synthesize new path

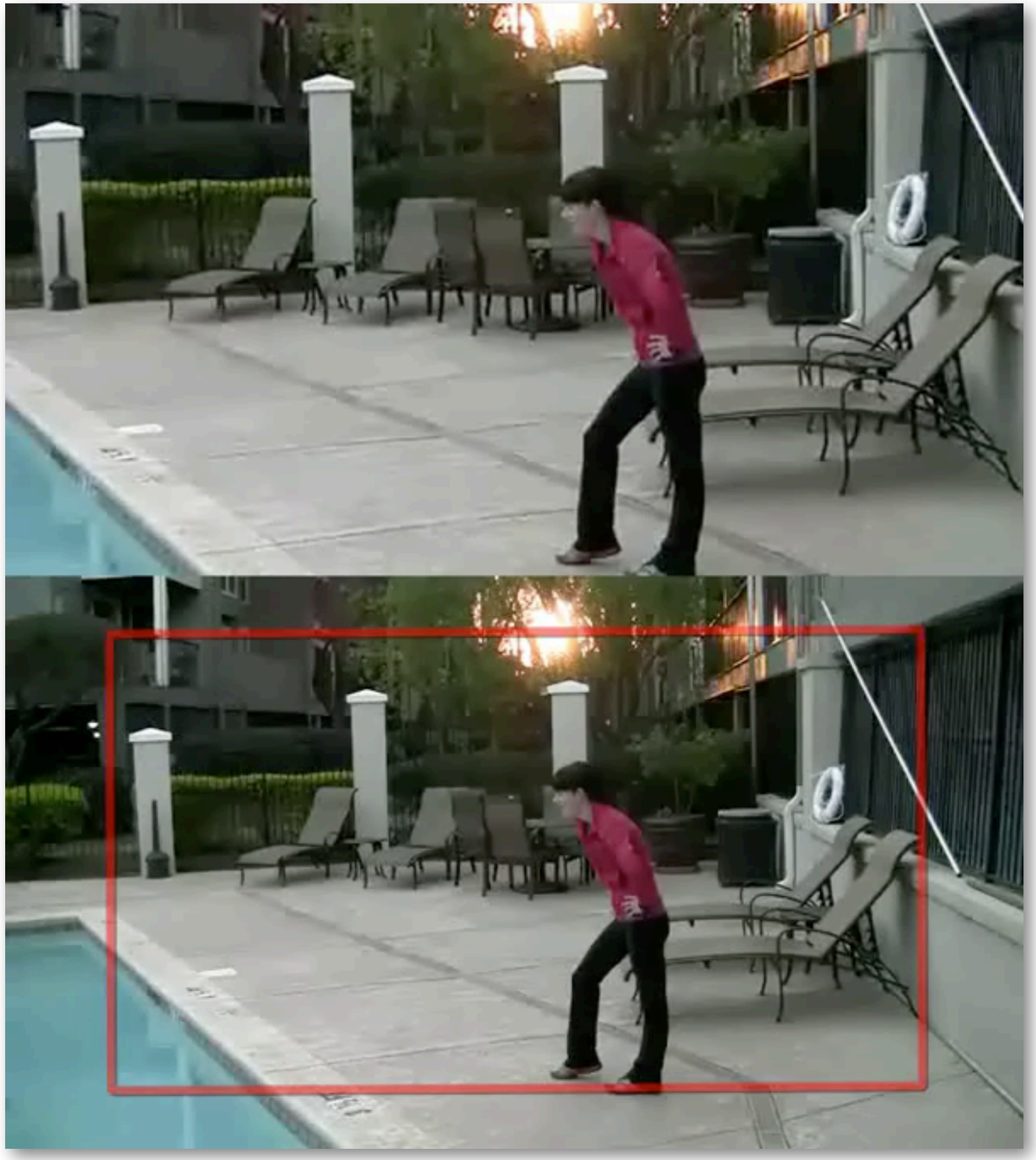
(1) Estimate Camera Motion

(2) Stabilize Camera Path

(3) Crop and Re-synthesize

Stabilization by Cropping

* Crop is constrained to stay within frame bounds (stable path within envelope)
* Apply virtual crop to yield stable video



YouTube Example

original (with crop)

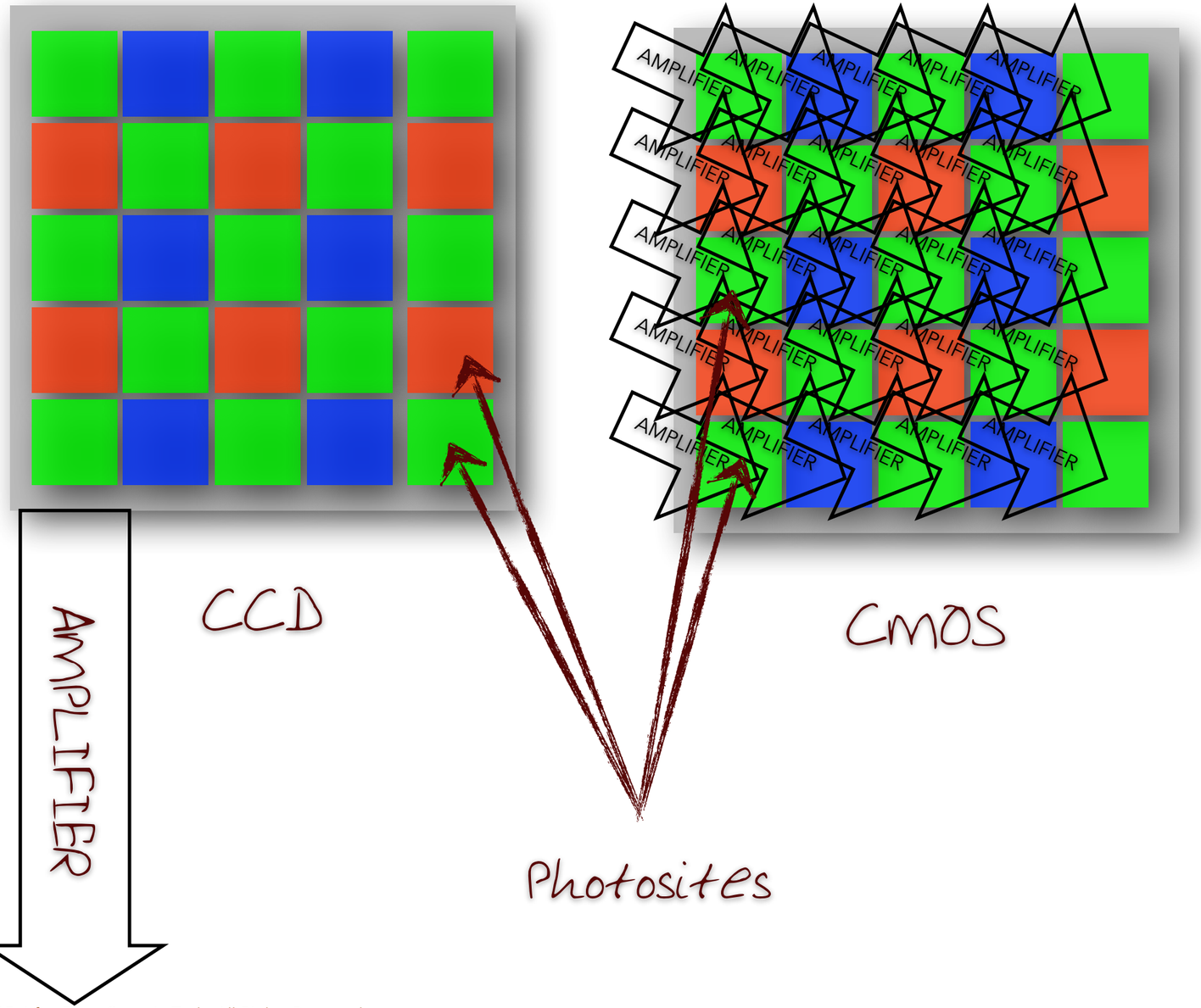
stabilized

YouTube Example

original

stabilized

Recall: CCD vs. CMOS Sensors



* CMOS: Complementary metal Oxide Semiconductor
* Photosites in CCD are passive and do no "work"
* Photosites in CMOS are amplifiers and can do local processing

Casual Videos with Rolling Shutter

Rolling Shutter Removal

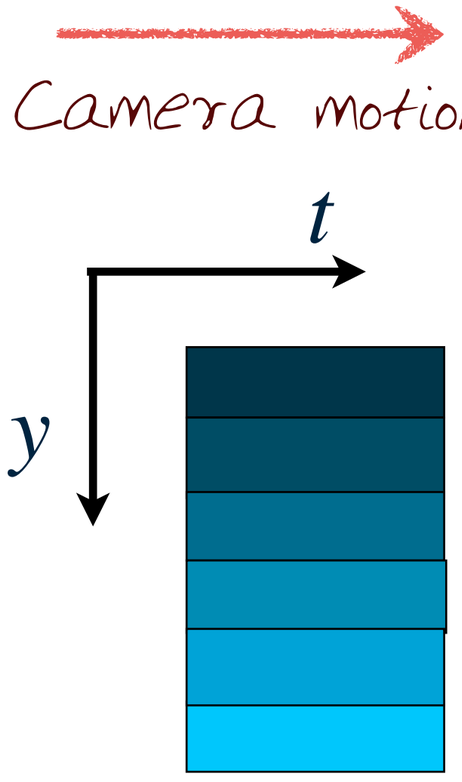
original (deliberately shaken)

rolling shutter removed

Types of electronic shutters

* Global shutter (CCD sensor)

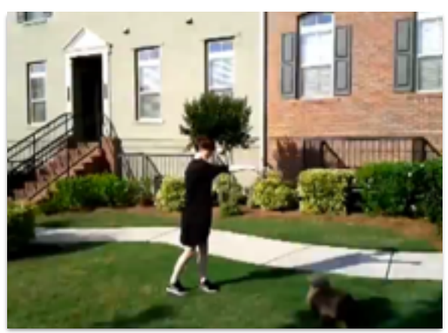
* Image read at one instant at time



Types of electronic shutters

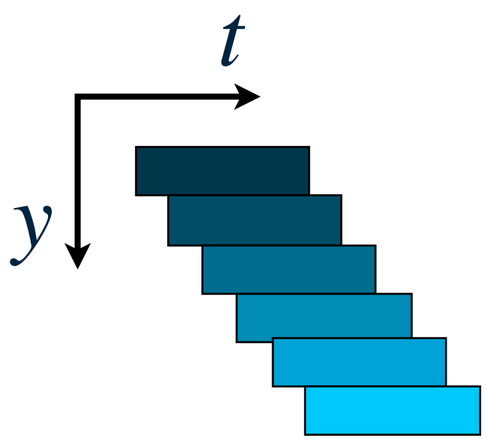
* Rolling shutter (CMOS sensor)
  * Image read one scan line at time

Camera motion



global shutter

rolling shutter



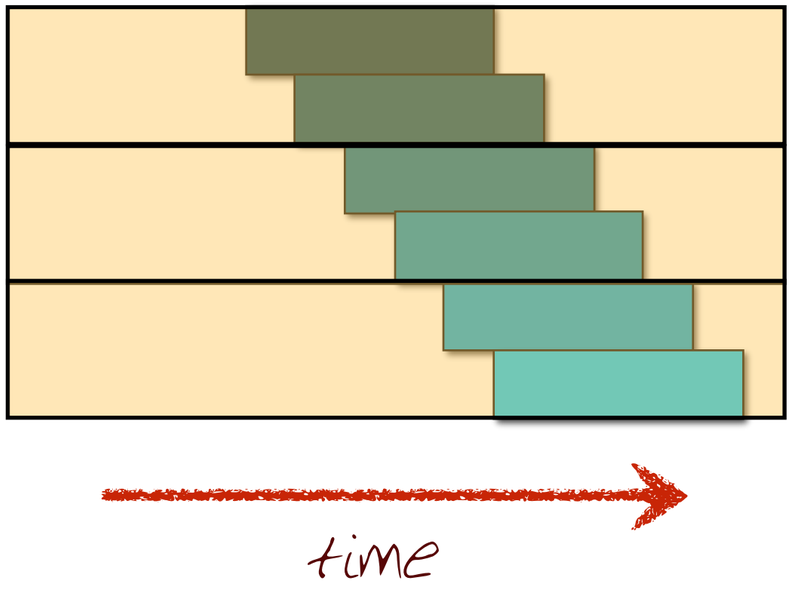
Global shutter model

Original courtesy of [Baker et al., 2010]

Global shutter model

Stabilized without rolling shutter removal [Youtube 2011]

Motion model mixtures



* Difficulty: Speed of readout varies across cameras
* Solution: Use multiple motion models and blend using mixtures of Gaussians

Rolling shutter model

Original courtesy of [Baker et al., 2010]

Rolling shutter model

Rolling shutter removed [YouTube 2012]

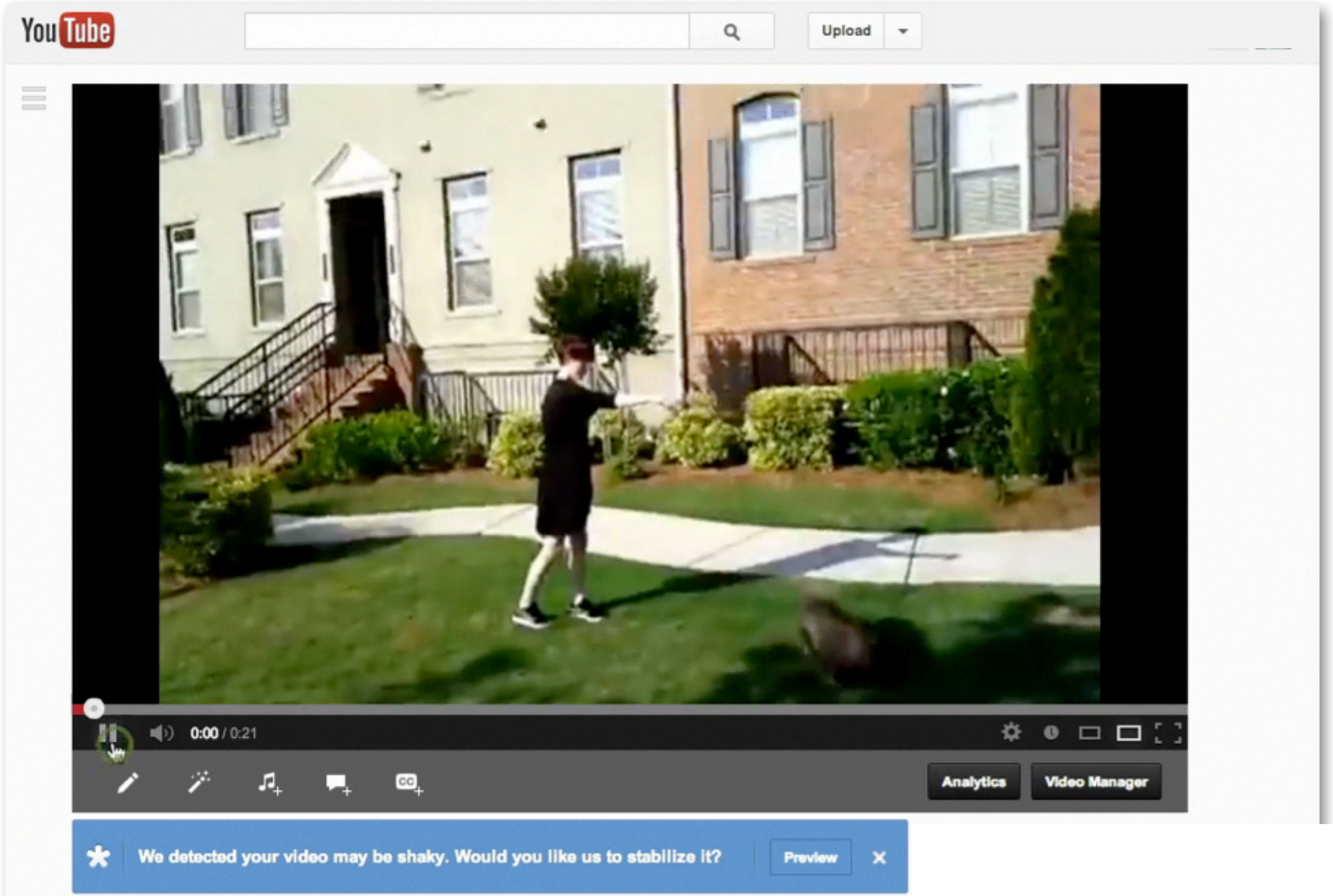
Rolling Shutter Wobble

Rolling Shutter Wobble

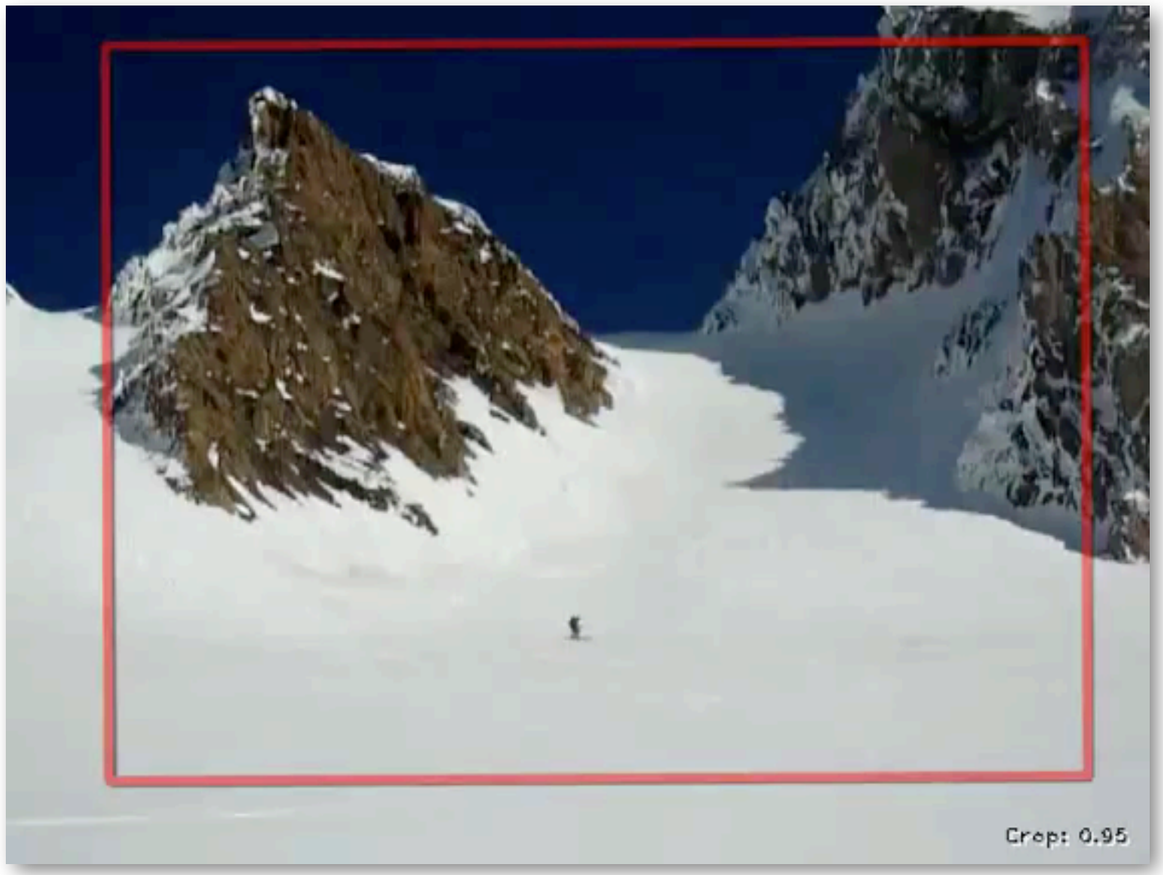
* Removed &
  Stabilized

original

stabilized



Adaptive Shake: Auto Crop



original with crop

stabilized result

Alternative Stabilizer

https://www.youtube.com/watch?v=_dPlkFPowCc

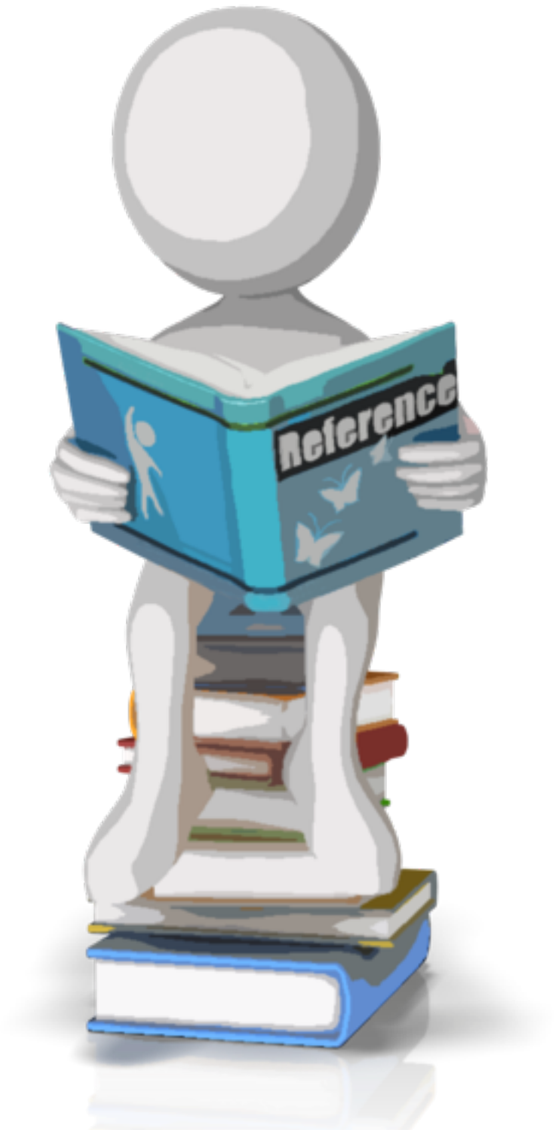
Ad for LG G2

https://www.youtube.com/watch?v=VRKu785g0O0

Summary

* Introduced Video Stabilization
* Demoed a working system
* Discussed DOF for camera motion modeling
* Discussed Rolling Shutter and how it can be removed for Video Stabilization

Further Reading



* Grundmann, Kwatra, & Essa (2011), "Auto-Directed Video Stabilization with Robust L1 Optimal Camera Paths," IEEE CVPR, 2011. [PDF]
* Grundmann, Kwatra, Castro, & Essa (2012), "Calibration-Free Rolling Shutter Removal," in IEEE ICCP, 2012. [PDF]
* Baker, Bennett, Kang, & Szeliski (2010) "Removing Rolling Shutter Wobble" in IEEE CVPR 2010 [PDF]
* Forssén, & Ringaby (2010) "Rectifying rolling shutter video from hand-held devices" in IEEE CVPR 2010 [PDF]

Credits

* For more information, see
  * M. Grundmann, V. Kwatra, and I. Essa (2011), "Auto-Directed Video Stabilization with Robust L1 Optimal Camera Paths," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. [PDF]
  * M. Grundmann, V. Kwatra, D. Castro, and I. Essa (2012), "Calibration-Free Rolling Shutter Removal," in Proceedings of IEEE Conference on Computational Photography (ICCP), 2012. WINNER OF BEST PAPER AWARD [PDF]
  * http://prof.irfanessa.com/tag/video-stabilization/

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing

Panoramic Video Textures

* Combine Video Textures and Panorama

Lesson Objectives

1. Review Video Textures and Panoramas
2. Combine Video Textures and Panoramas to form Panoramic Video Textures
3. Construct a Panorama from Video
4. Construct a Video Texture from Dynamic parts of the scene

Panoramic Video Textures

* A panoramic video texture (PVT)
* A video that has been stitched into a single, wide field of view
* Appears to play continuously and indefinitely

Agarwala, Zheng, Pal, Agrawala, Cohen, Curless, Salesin, Szeliski (2005) "Panoramic Video Textures" SIGGRAPH 2005

Recall: Panoramas

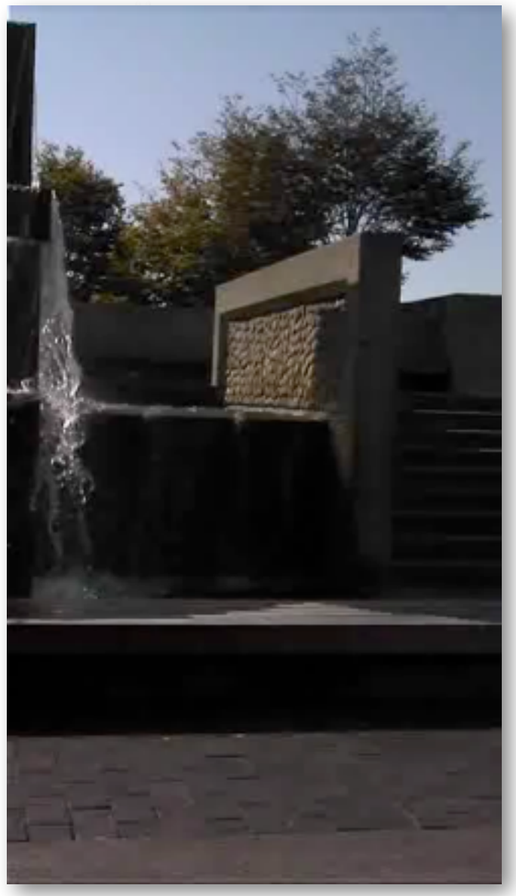
Recall: Video Texture

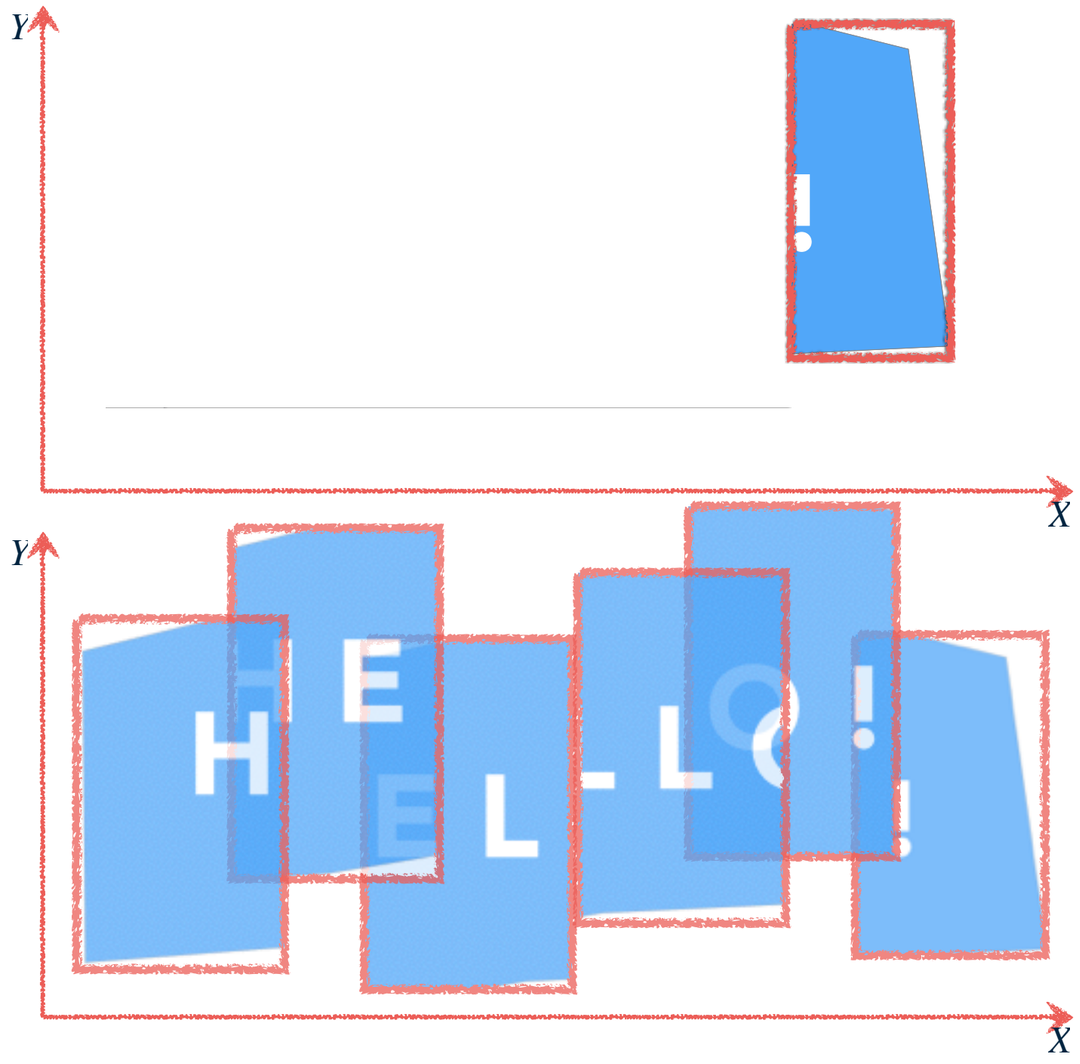
(Schödl, Salesin, Szelsiki, Essa 2003)

Panoramic Video Textures

QUIZ

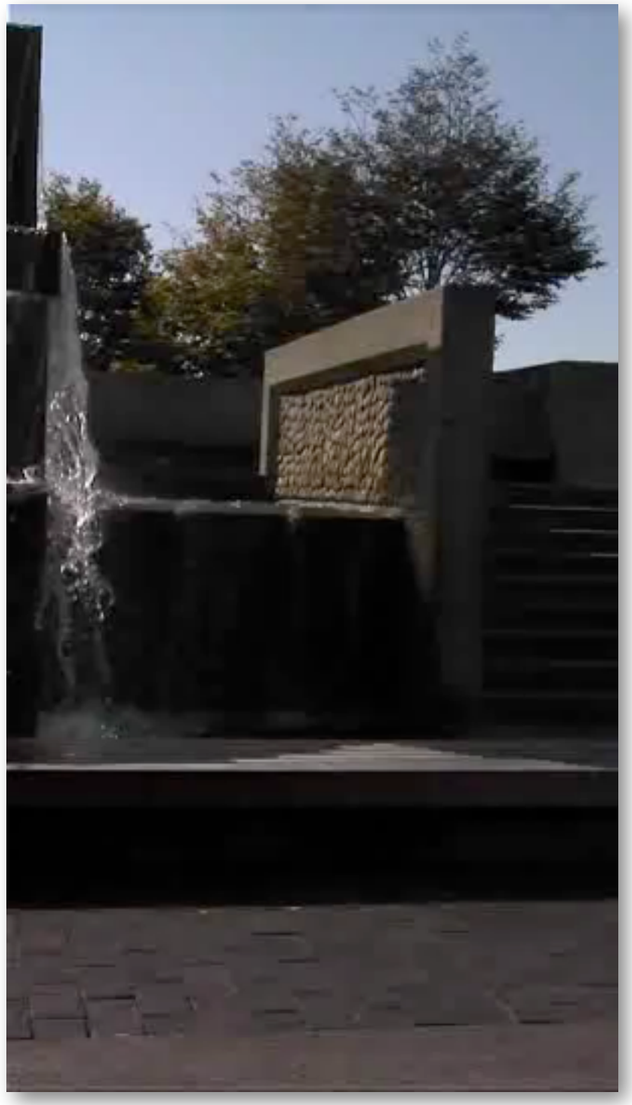
Video Pan to a Panorama?

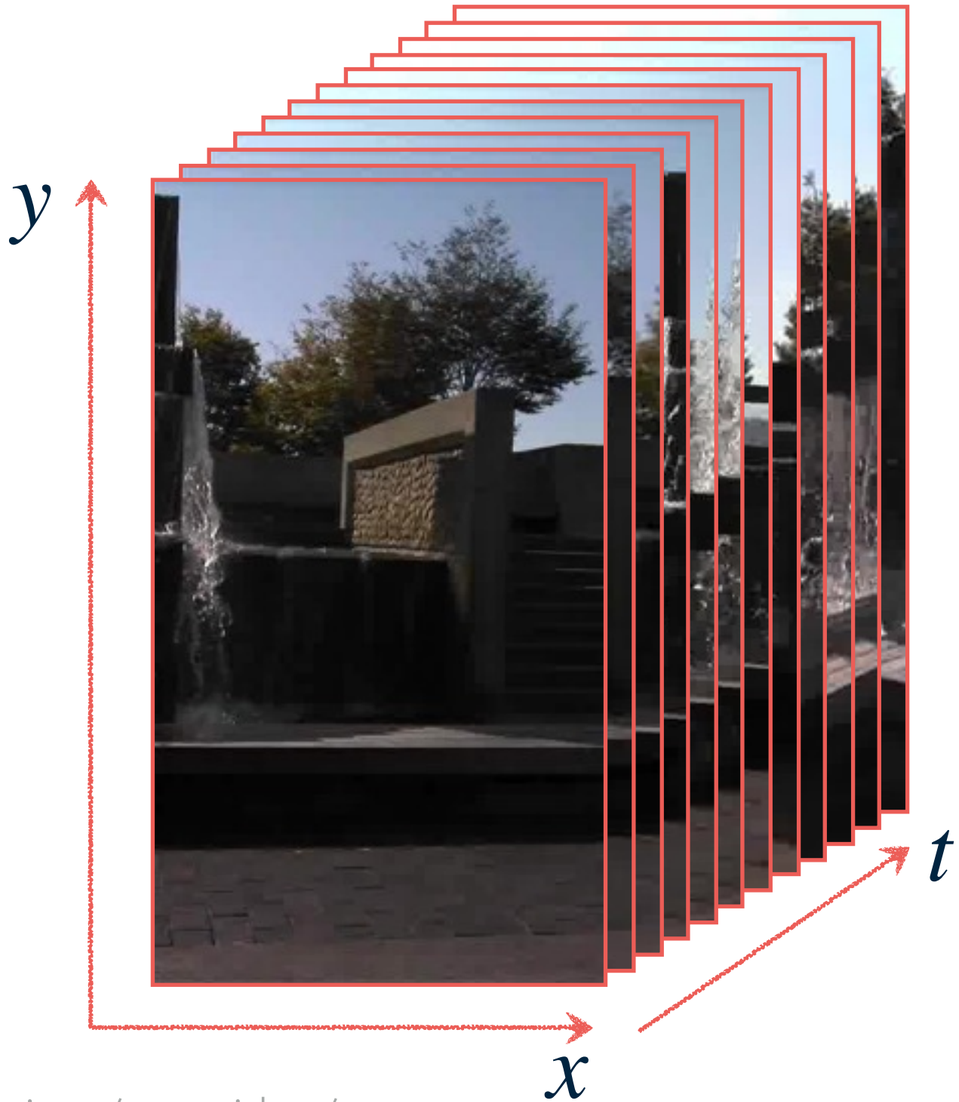




- ☐ Use all the video frame, register and align to generate panorama
- ☐ Take select 6 images and generate panorama
- ☐ Both

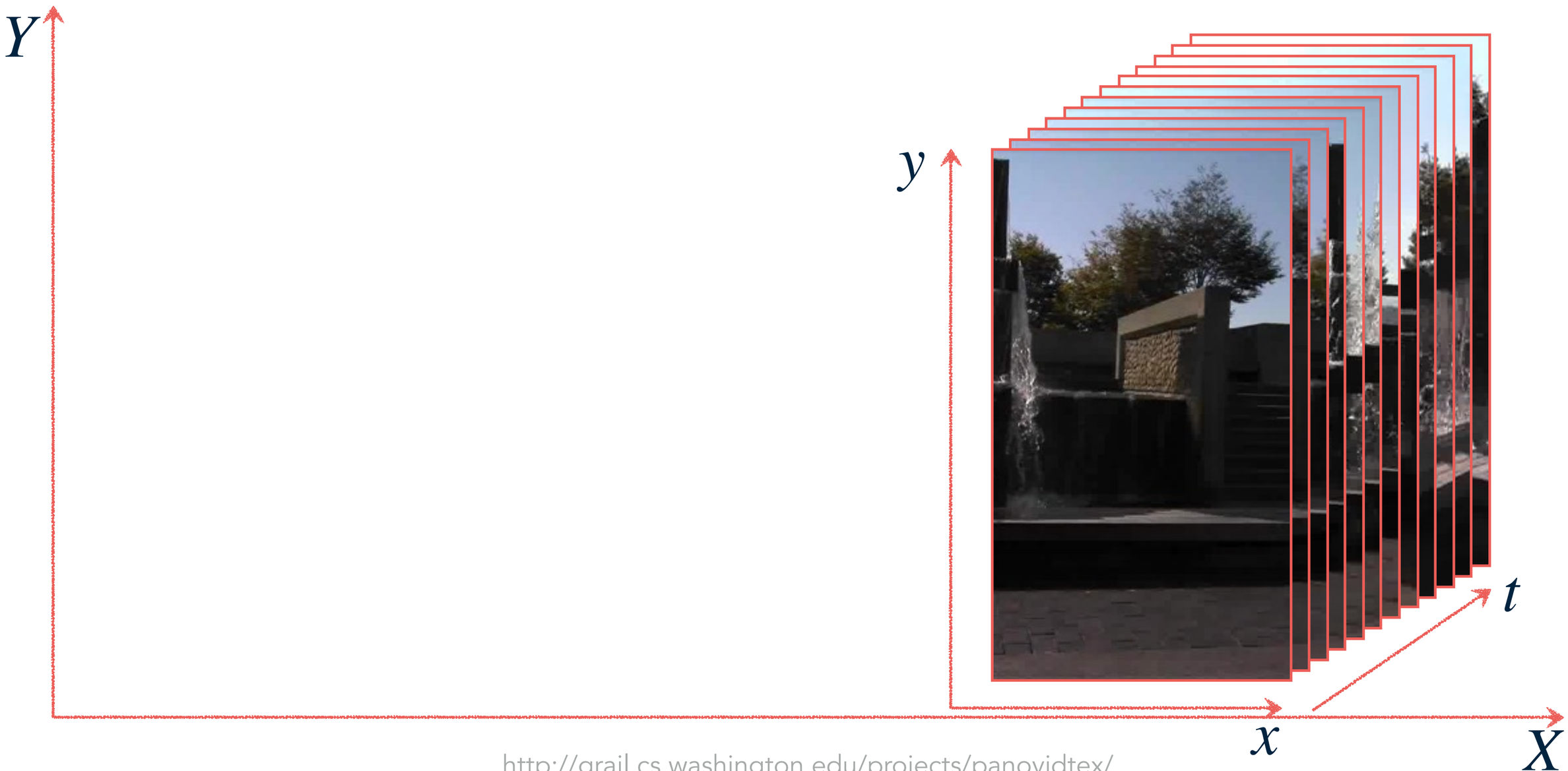
Video Registration





http://grail.cs.washington.edu/projects/panovidtex/

Video Registration



http://grail.cs.washington.edu/projects/panovidtex/

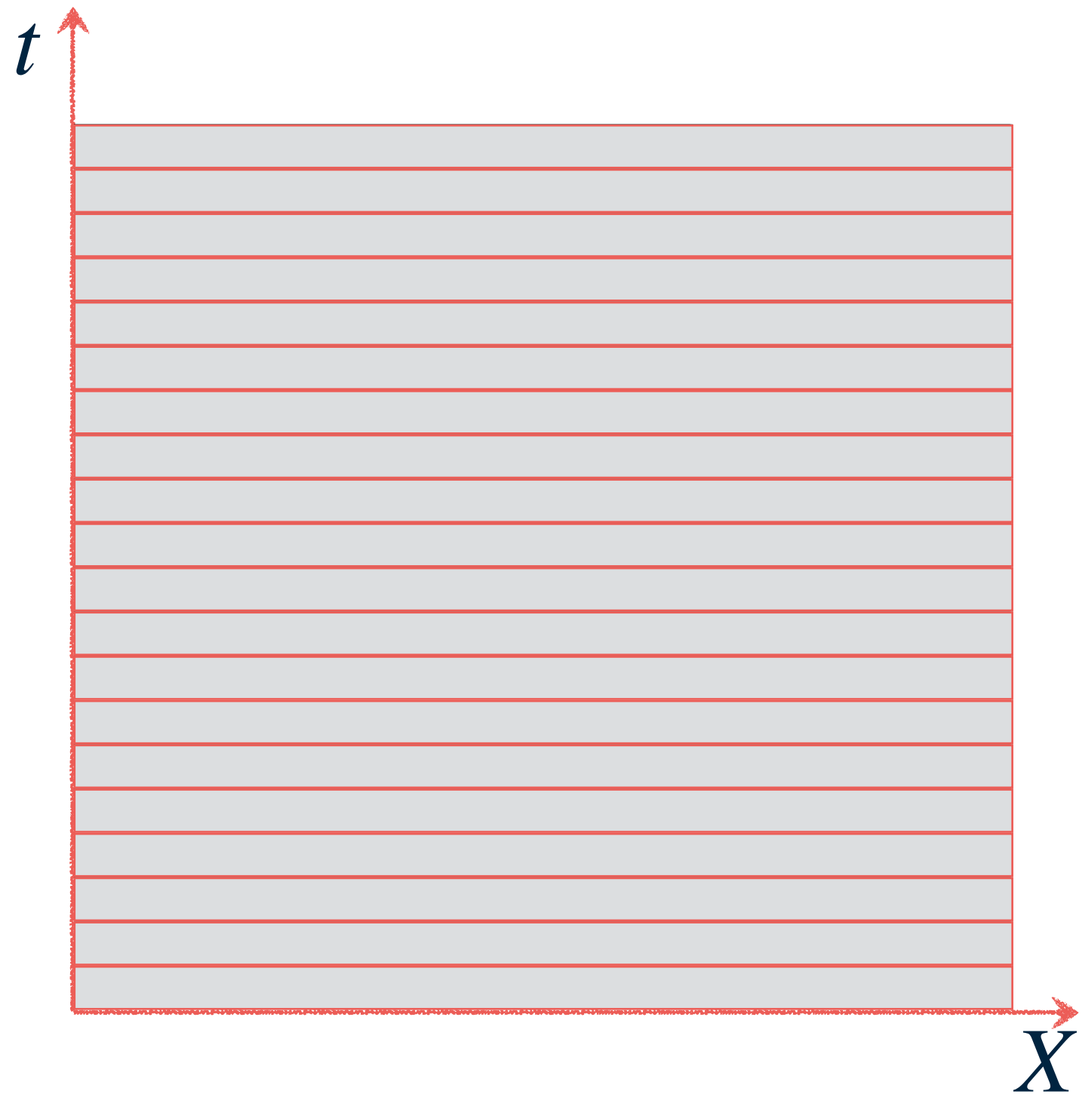
Video Registration

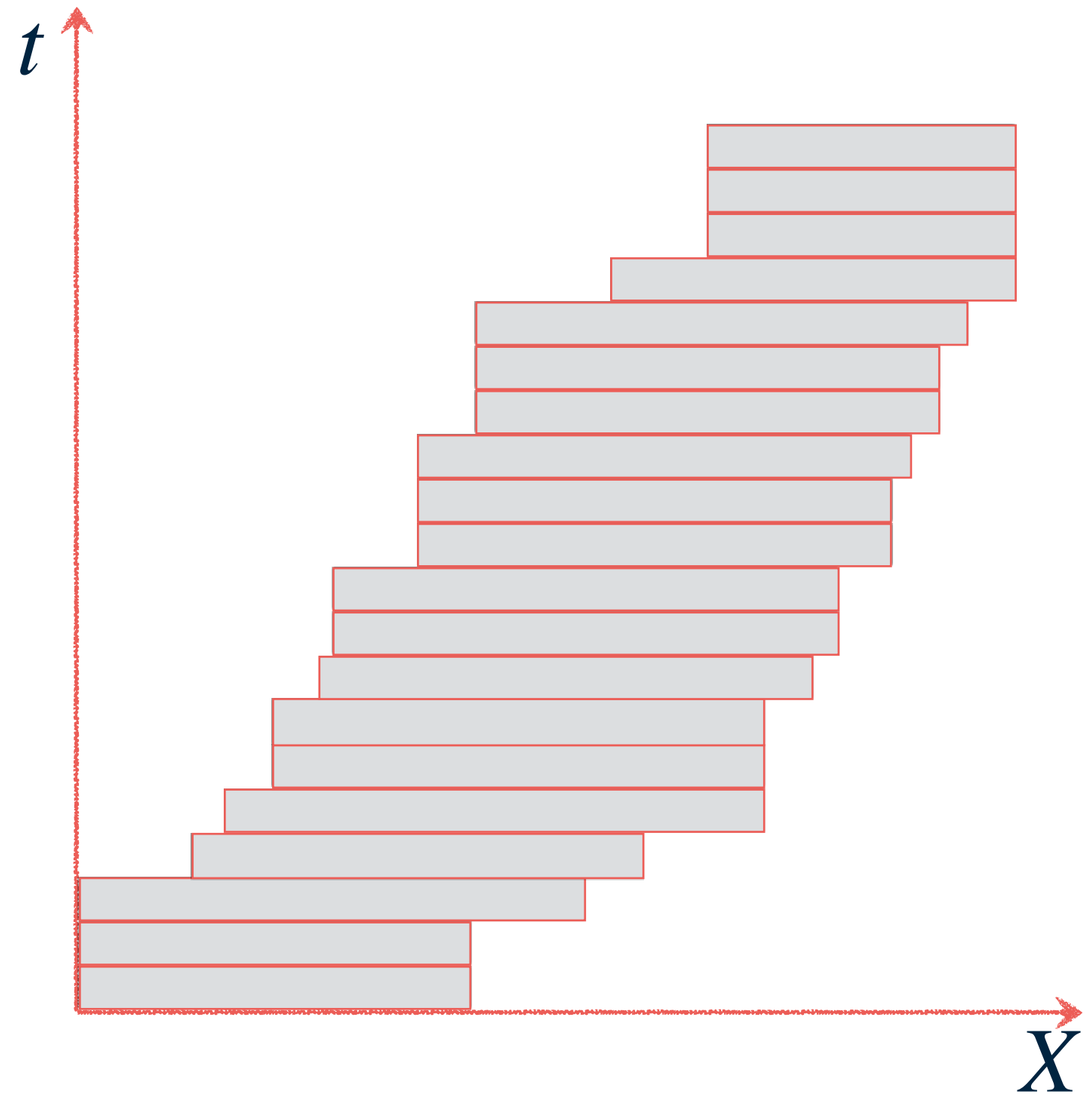
http://grail.cs.washington.edu/projects/panovidtex/

Static vs. Dynamic Regions

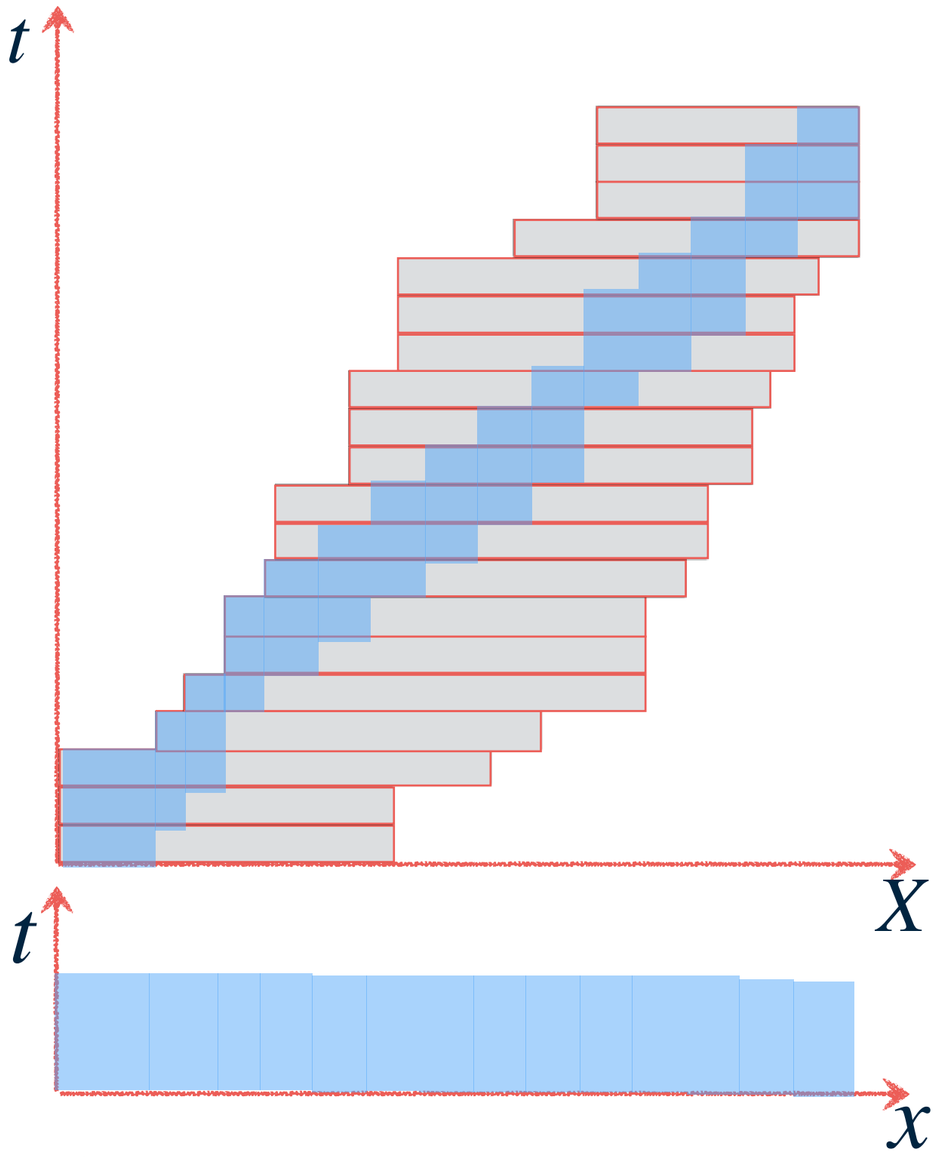
http://grail.cs.washington.edu/projects/panovidtex/

Video Textures of Dynamic Regions





Video Textures of Dynamic Regions



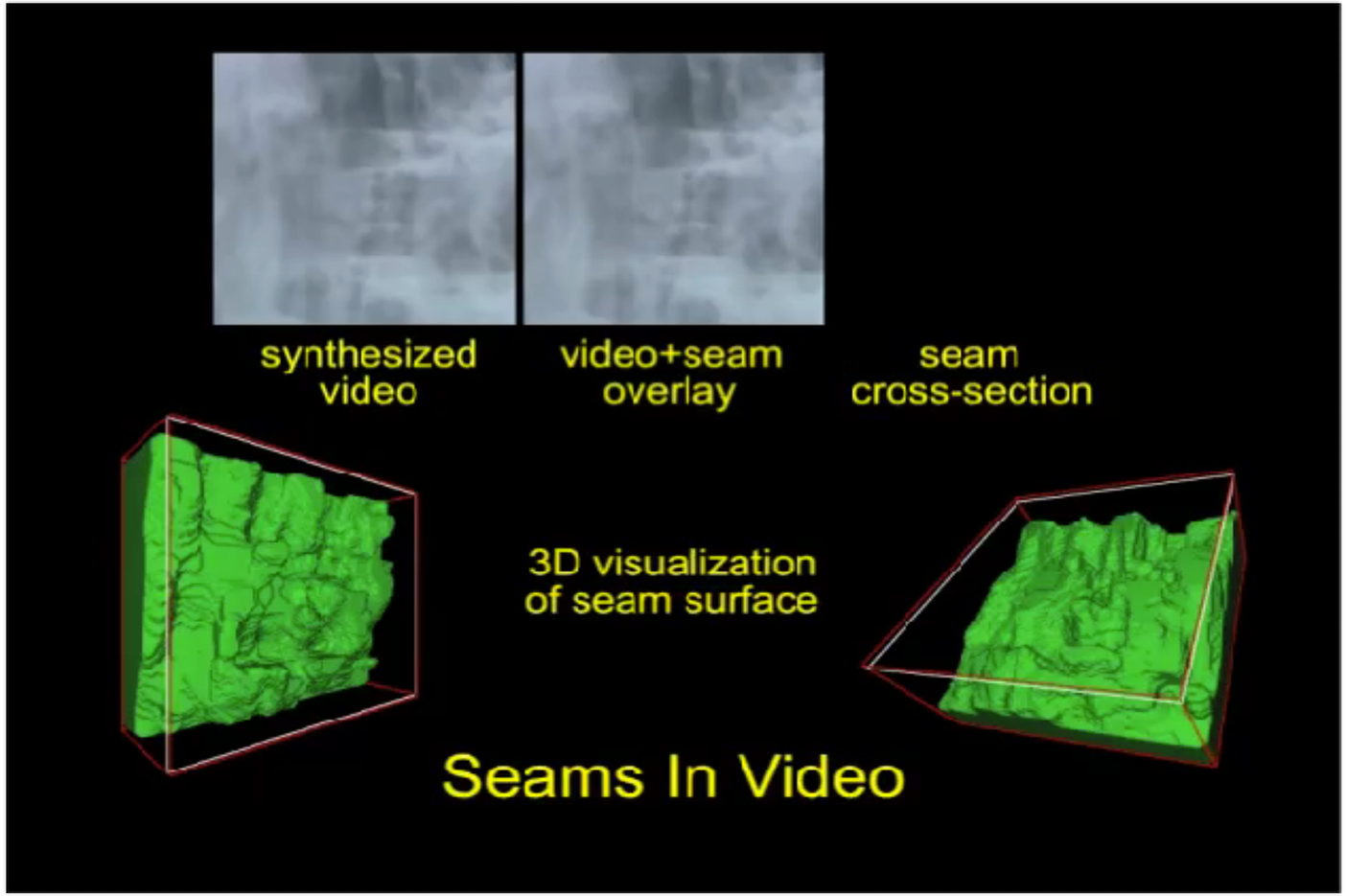
* map a continuous diagonal slice of the input video volume to the output panorama
* Restricts boundaries to frames
* Shears spatial structures across time

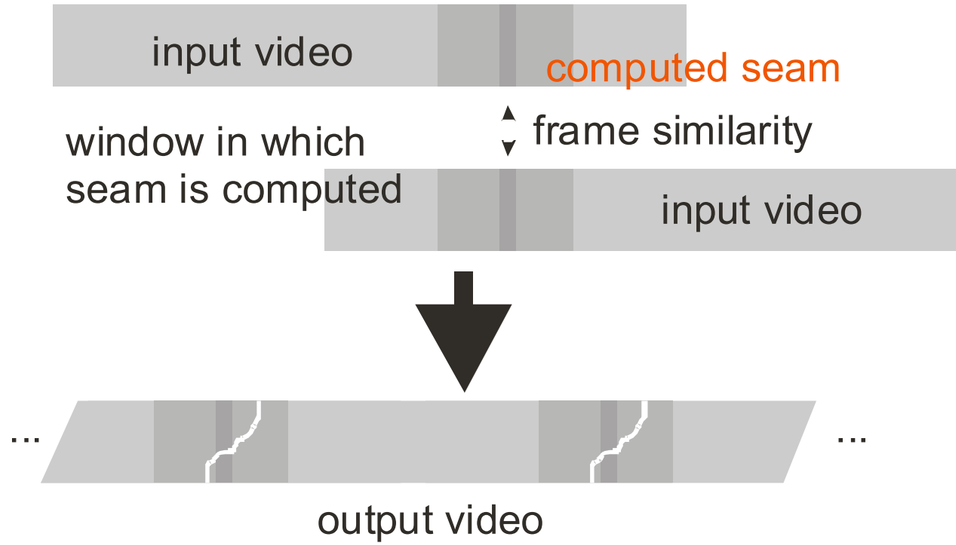
Video Textures of Dynamic Regions

Original

Continuous Diagonal Slices

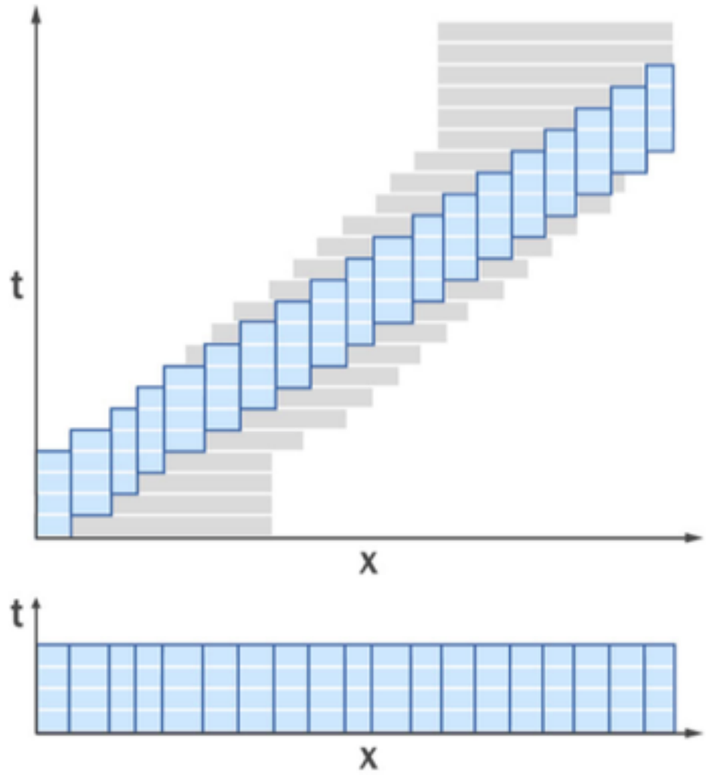
Not Just Fade/Blend, but Cut



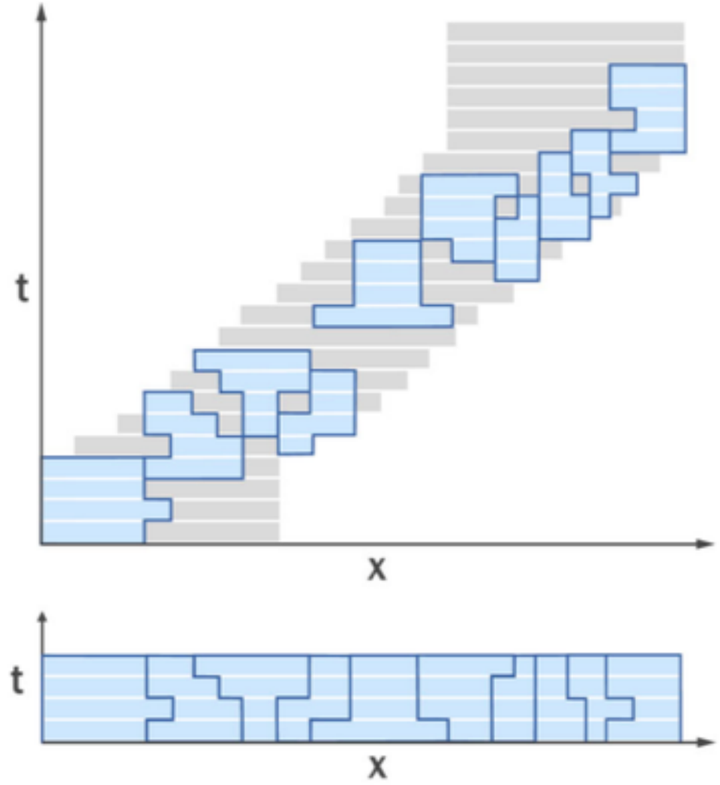


(Kwatra, Schödl, Essa 2003)

Video Textures of Dynamic Regions



map a continuous diagonal slice of the input video volume



mapped to locations in the input in coherent fragments

Agarwala et al. (2005)

Video Textures of Dynamic Regions

Original

Coherent Fragments

Summary

* Introduced the concept of Panoramic Video Textures
* Extension of Video Textures and Panoramas

Further Reading

* Agarwala, Zheng, Pal, Agrawala, Cohen, Curless, Salesin, Szeliski (2005) "Panoramic Video Textures" SIGGRAPH 2005 [PDF]
* Schödl, Szeliski, Salesin, and Essa (2000), "Video textures," in SIGGRAPH 2000 [PDF]
* Kwatra, Schödl, Essa, Turk, Bobick (2003), "Graphcut textures: image and video synthesis using graph cuts" in SIGGRAPH 2003 [PDF]

Credits

* For more information, see
  * Agarwala, Zheng, Pal, Agrawala, Cohen, Curless, Salesin, Szeliski (2005) "Panoramic Video Textures" SIGGRAPH 2005
* http://grail.cs.washington.edu/projects/panovidtex/

Computational Photography

* Study the basics of computation and its impact on the entire workflow of photography, from capturing, manipulating and collaborating on, and sharing photographs.

Georgia Tech College of Computing