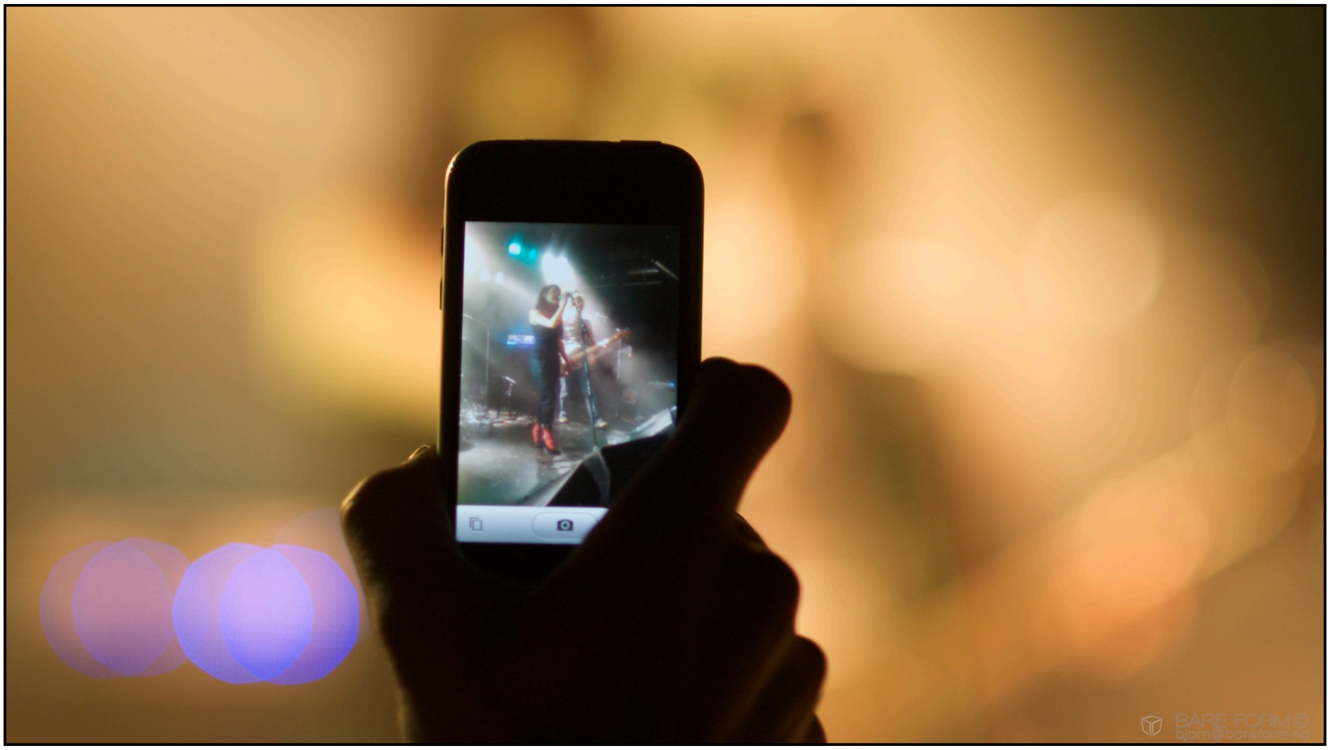
Introduction and Feature Detection

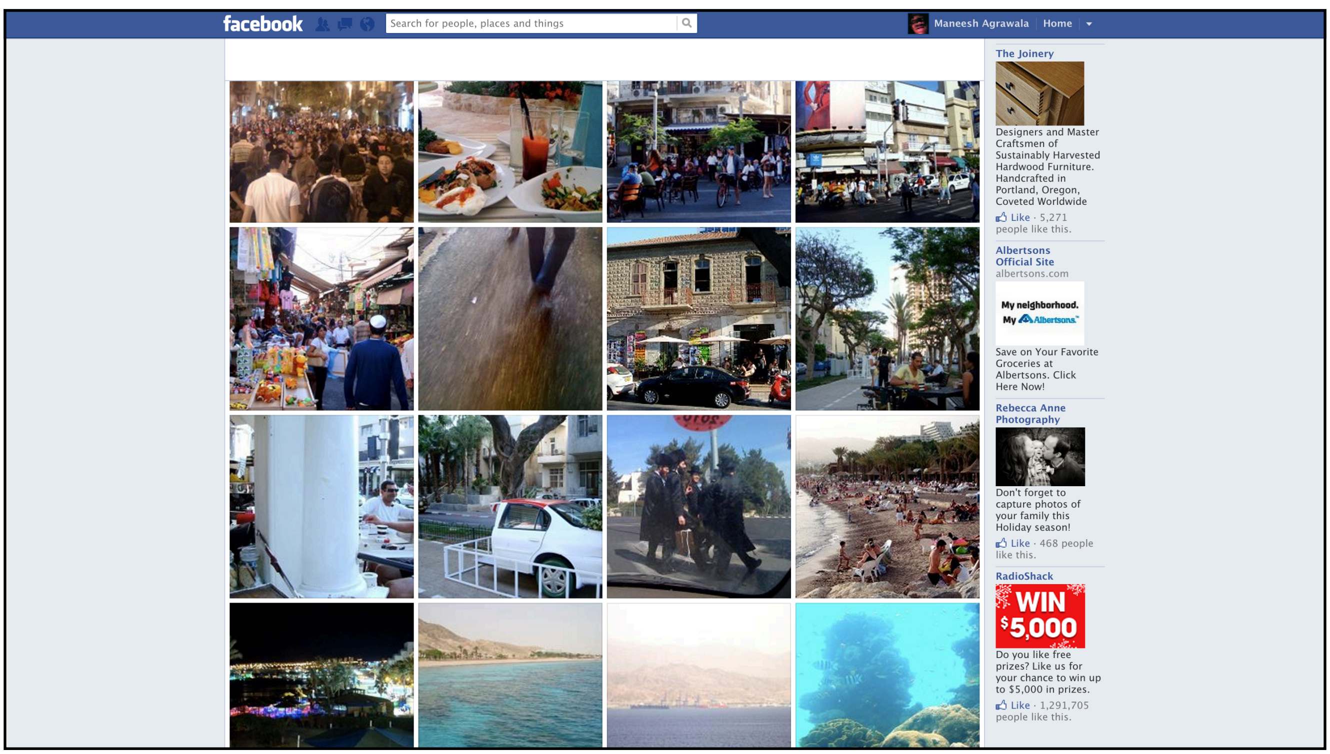
CS448V — Computational Video Manipulation

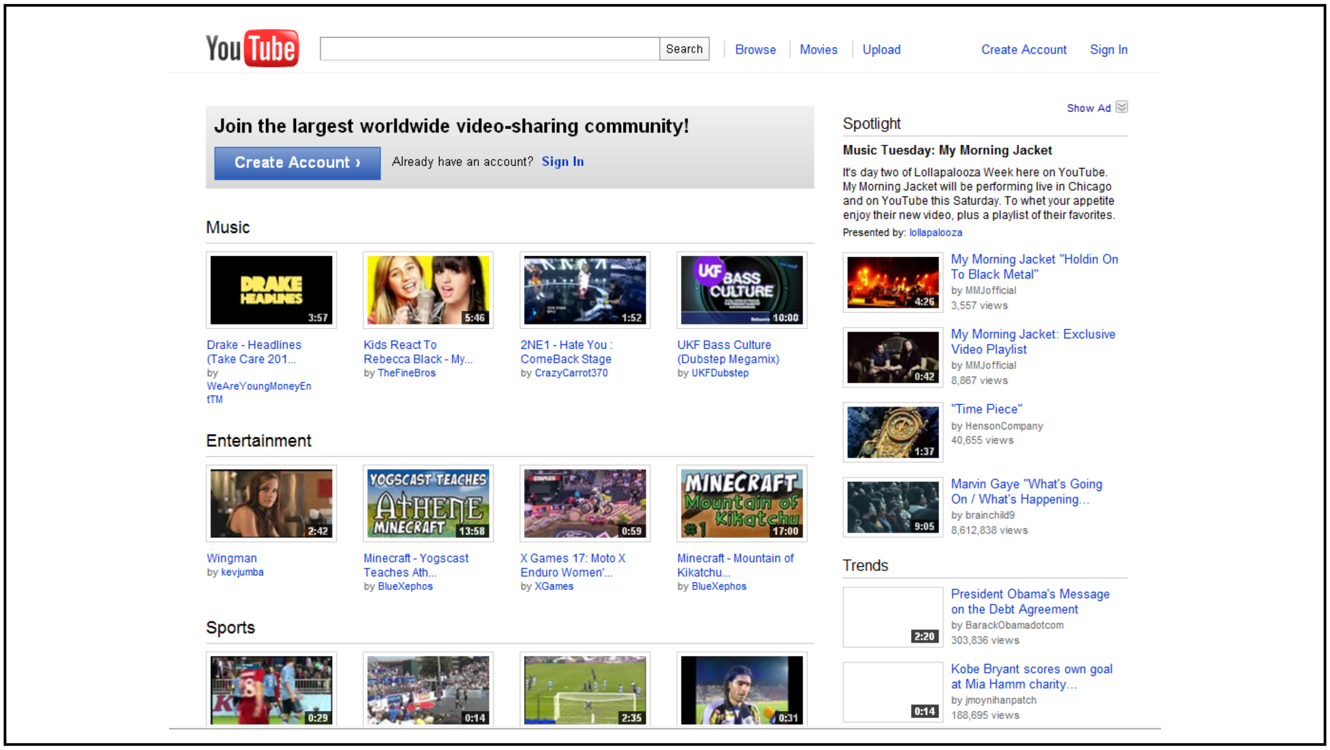
April 2019

Raiders of the Lost Ark: The Adaptation [Zala 82-89]
Shot-for-shot remake by three 12-year olds (took 7 years)

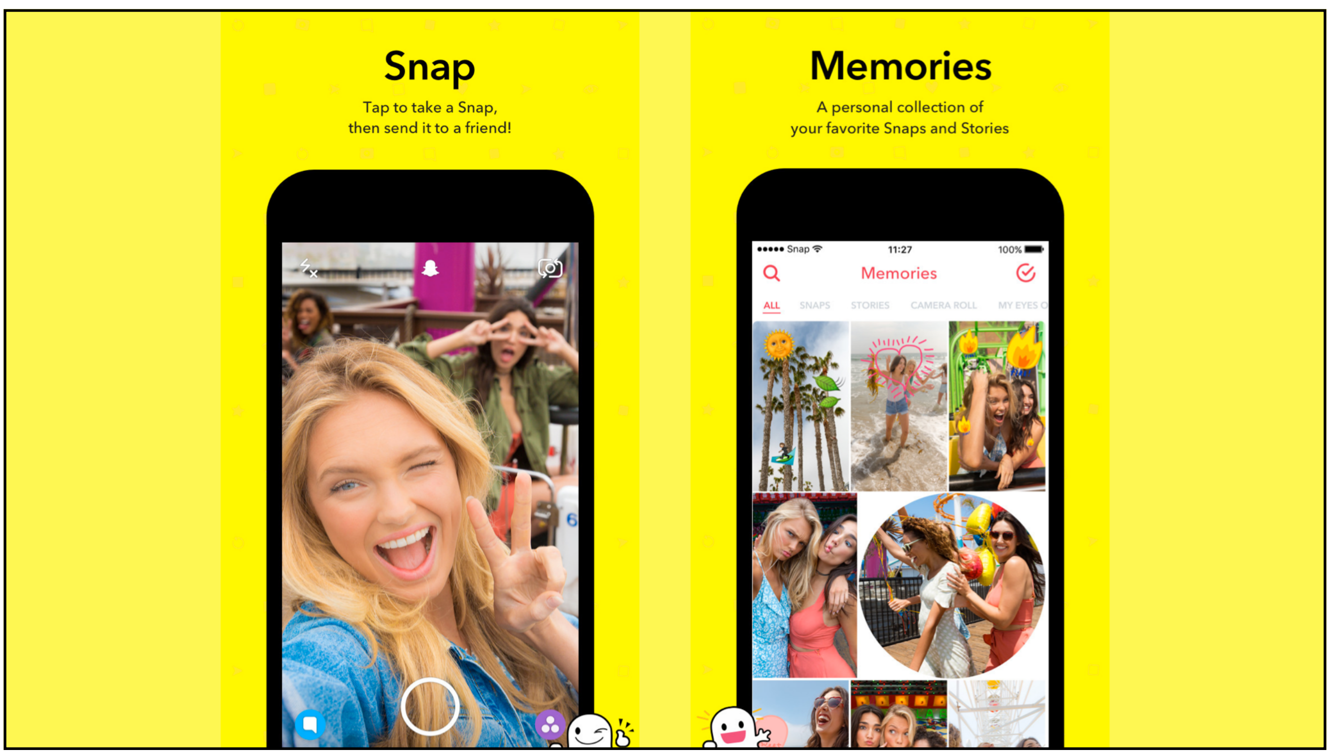
How can we let people easily make such video?



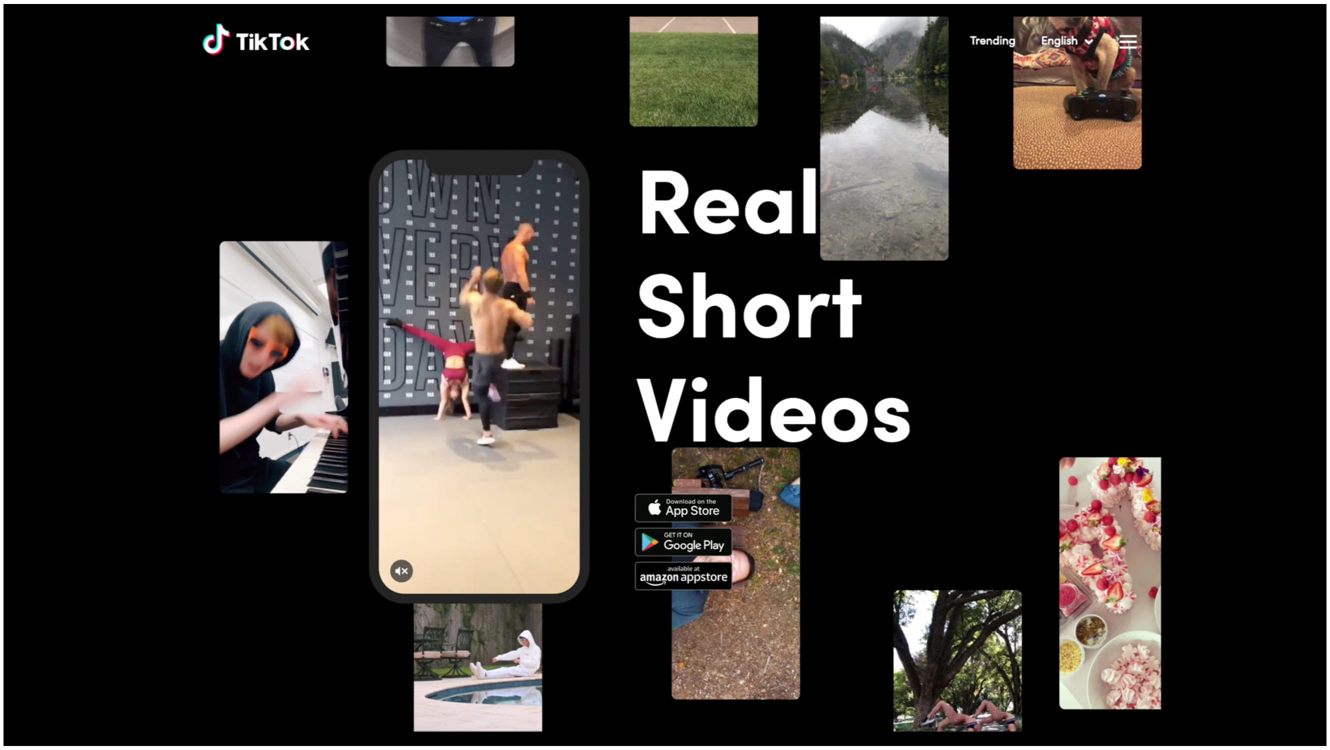

facebook
Search for people, places and things
Maneesh Agrawala
Home
The Joinery
Designers and Master Craftsmen of Sustainably Harvested Hardwood Furniture. Handcrafted in Portland, Oregon, Coveted Worldwide
Like · 5,271 people like this.
Albertsons Official Site
albertsons.com
My neighborhood. My Albertsons
Save on Your Favorite Groceries at Albertsons. Click Here Now!
Rebecca Anne Photography
Don't forget to capture photos of your family this Holiday season!
Like · 468 people like this.
RadioShack
WIN $5,000
Do you like free prizes? Like us for your chance to win up to $5,000 in prizes.
Like · 1,291,705 people like this.



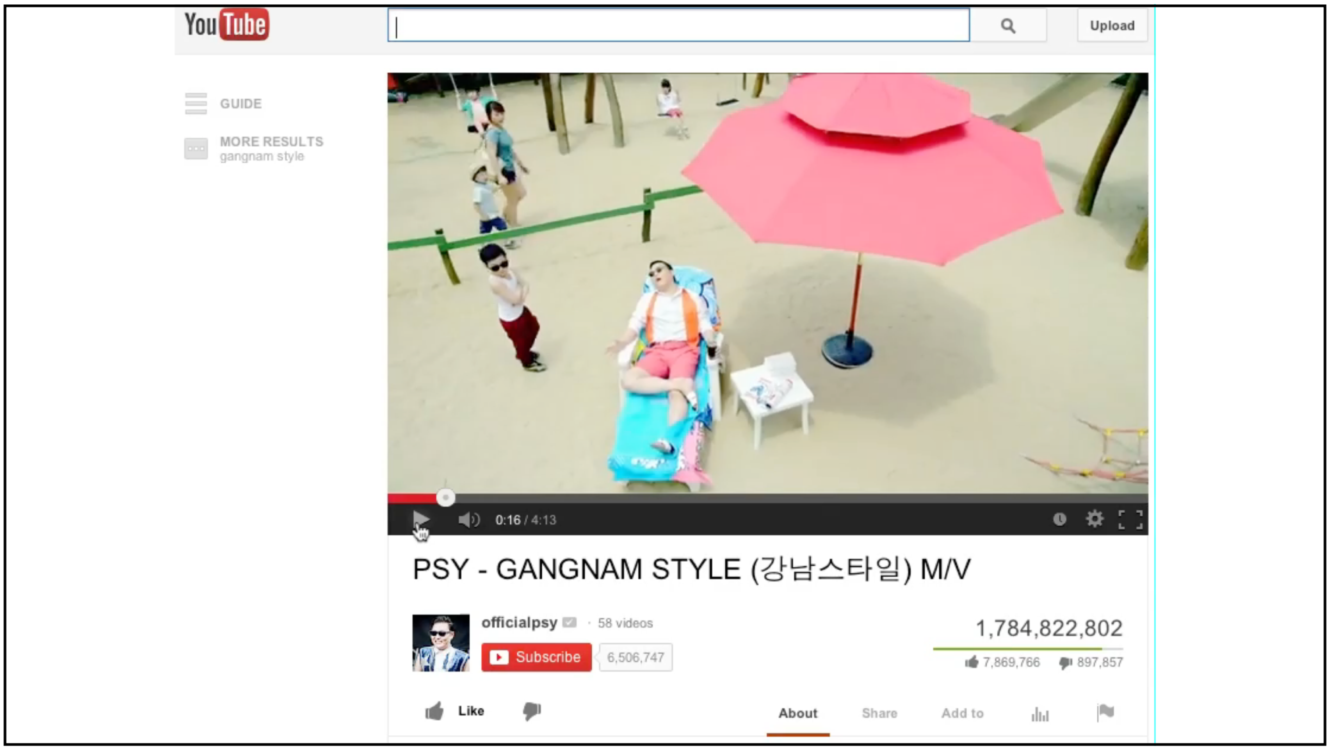
YouTube
Search
Browse
Movies
Upload
Create Account
Sign In
Join the largest worldwide video-sharing community!
Create Account ›
Already have an account? Sign In
Music
Drake - Headlines (Take Care 201...
by WeAreYoungMoneyEntTM
Kids React To Rebecca Black - My...
by TheFineBros
2NE1 - Hate You : ComeBack Stage
by CrazyCarrot370
UKF Bass Culture (Dubstep Megamix)
by UKFDubstep
Entertainment
Wingman
by kevjumba
Minecraft - Yogscast Teaches Ath...
by BlueXephos
X Games 17: Moto X Enduro Women'...
by XGames
Minecraft - Mountain of Kikatchu...
by BlueXephos
Sports
Show Ad
Spotlight
Music Tuesday: My Morning Jacket
It's day two of Lollapalooza Week here on YouTube. My Morning Jacket will be performing live in Chicago and on YouTube this Saturday. To whet your appetite enjoy their new video, plus a playlist of their favorites.
Presented by: lollapalooza
My Morning Jacket "Holdin On To Black Metal"
by MMJofficial
3,557 views
My Morning Jacket: Exclusive Video Playlist
by MMJofficial
8,867 views
"Time Piece"
by HensonCompany
40,655 views
Marvin Gaye "What's Going On / What's Happening...
by brainchild9
8,612,838 views
Trends
President Obama's Message on the Debt Agreement
by BarackObamadotcom
303,836 views
Kobe Bryant scores own goal at Mia Hamm charity...
by jmoynihanpatch
188,695 views

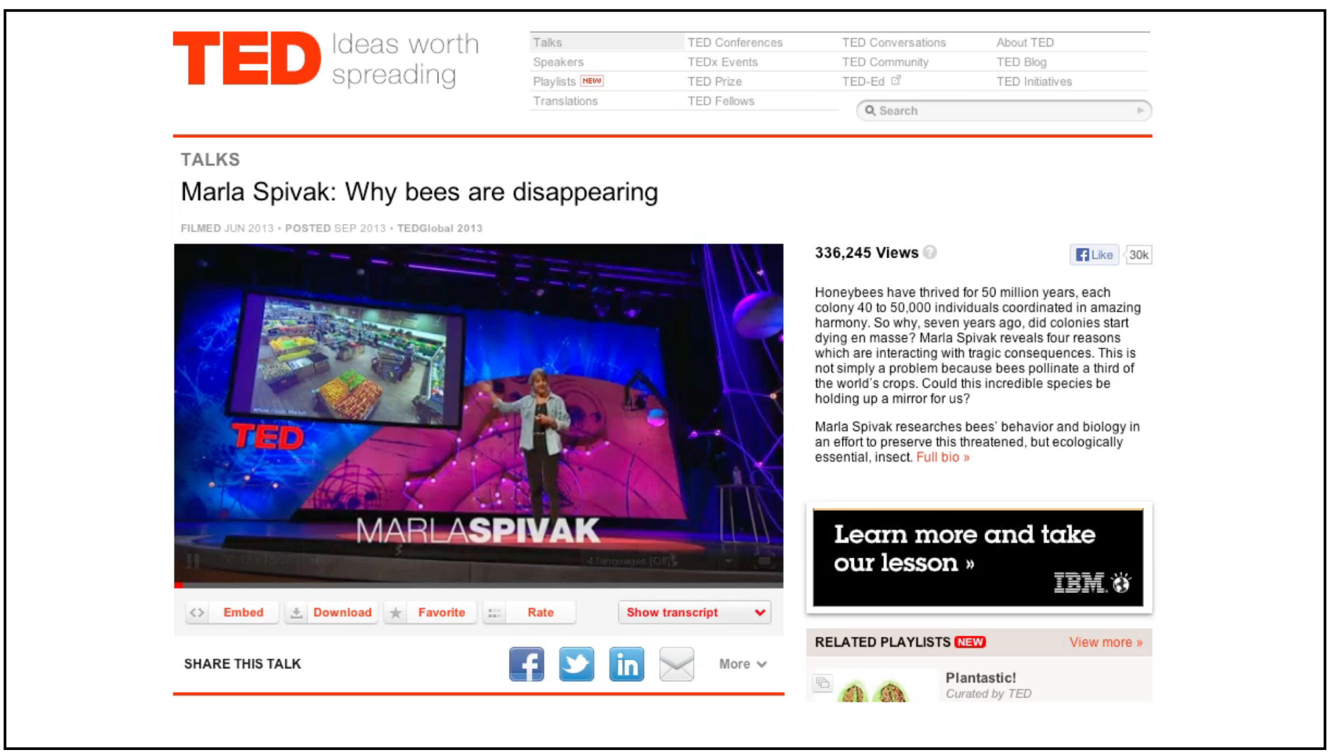


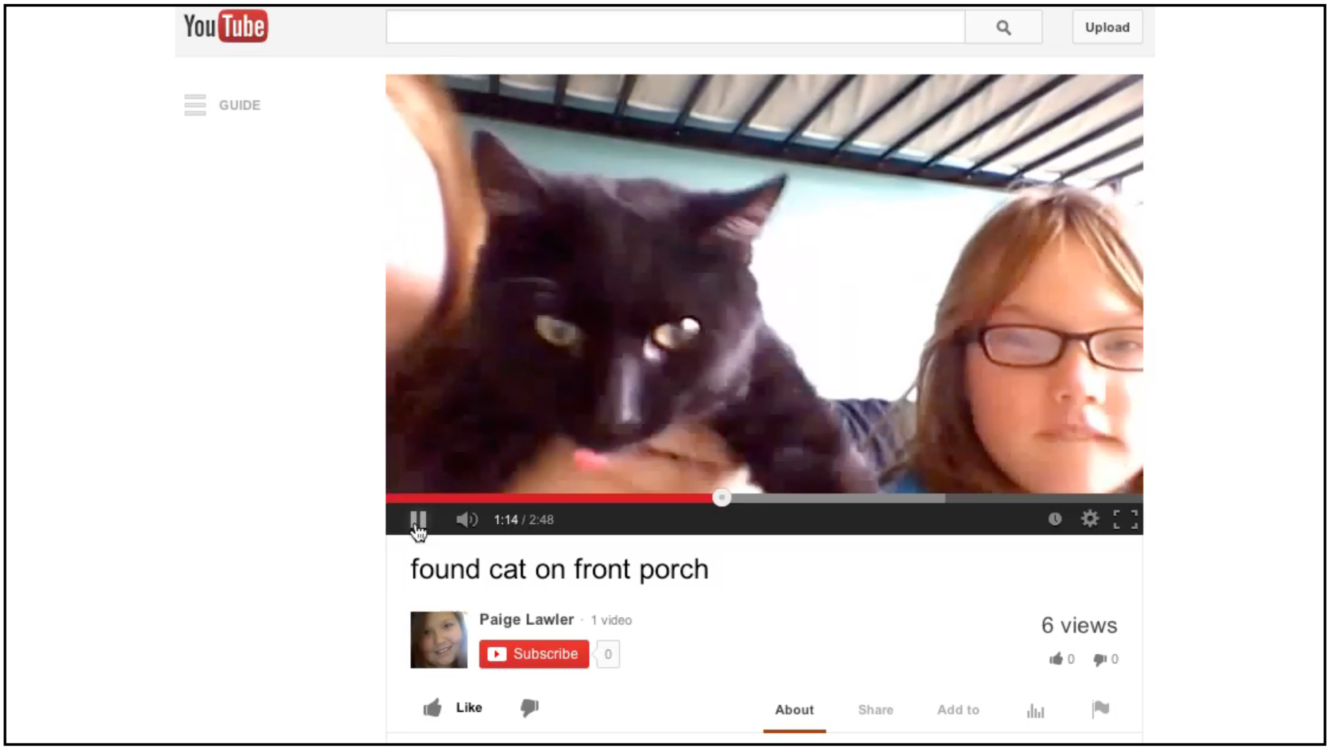
Snap
Tap to take a Snap, then send it to a friend!
Memories
A personal collection of your favorite Snaps and Stories
11:27
100%
Memories
ALL
SNAPS
STORIES
CAMERA ROLL


TikTok
Trending
English
Real
Short
Videos
Download on the App Store
GET IT ON Google Play
available at amazon appstore


YouTube
Search
鶏攻撃の術
0:00 / 3:06
Chicken Attack
2,218,588 views
50K
1.6K
SHARE
SAVE
Up next
AUTOPLAY
Chicken Attack Remix [official] (not clickbait)
schmoyoho
1M views
Chicken Attack - Takeo Ischi [Piano Synthesia]
Hung Phạm Vũ
128K views
The Ending of The Matrix Trilogy Finally Explained
Looper
Recommended for you
15 Incredibly RARE and UNIQUE firearms OF ALL TIME
Amazing Stock
Recommended for you
More adventures in replying to spam | James Veitch
TED Archive
Recommended for you
Funniest Game Show Answers of All Time
theSeanGshow
39M views
George Carlin on some cultural issues.
Anders Jacobsson
7.9M views


YouTube
Upload
GUIDE
MORE RESULTS
gangnam style
0:16 / 4:13
PSY - GANGNAM STYLE (강남스타일) M/V
officialpsy · 58 videos
Subscribe 6,506,747
1,784,822,802
7,869,766 897,857
Like
About Share Add to


TED Ideas worth spreading
Talks TED Conferences TED Conversations About TED
Speakers TEDx Events TED Community TED Blog
Playlists NEW TED Prize TED-Ed TED Initiatives
Translations TED Fellows
Search
TALKS
Marla Spivak: Why bees are disappearing
FILMED JUN 2013 • POSTED SEP 2013 • TEDGlobal 2013
336,245 Views
Like 30k
Honeybees have thrived for 50 million years, each colony 40 to 50,000 individuals coordinated in amazing harmony. So why, seven years ago, did colonies start dying en masse? Marla Spivak reveals four reasons which are interacting with tragic consequences. This is not simply a problem because bees pollinate a third of the world's crops. Could this incredible species be holding up a mirror for us?
Marla Spivak researches bees' behavior and biology in an effort to preserve this threatened, but ecologically essential, insect. Full bio »
MARLASPIVAK
Learn more and take our lesson »
IBM
Embed Download Favorite Rate Show transcript
SHARE THIS TALK
More
RELATED PLAYLISTS NEW View more »
Plantastic!
Curated by TED


You Tube
Upload
GUIDE
1:14 / 2:48
found cat on front porch
Paige Lawler · 1 video
Subscribe
0
6 views
Like
About
Share
Add to


Adobe Premiere Pro - E:\Story_OnLo_PPro\Premiere Rev Workshop.prproj *
File Edit Project Clip Sequence Marker Title Window Help
Workspace: Editing
Access CS Live
Project: Premiere Rev Workshop
Source: A009_C003_0314G7_001.R3D 206-3
Effect Controls
Audio Mix
Program: !Medieval DVCPro Sequence
02:13:22:21
00:01:44:19
00:00:32:10
00:01:09:14
!Medieval DVCPro Sequence
Video 4
Video 3
Video 2
Video 1
Prod Sound 1

Challenge

People want to create and share stories

26% of all Internet users post original videos [Pew 13]
3,500,000,000 snaps/day uploaded to Snap [The Verge 17]
300 hours video/minute uploaded to YouTube [Youtube FAQ 18]

But raw video rarely tells a compelling story

Content not well thought out
Poor composition, lighting, etc.
Often too long

Best stories are planned, edited and produced

Current tools force users to work with low-level controls

Need higher-level tools for manipulating video

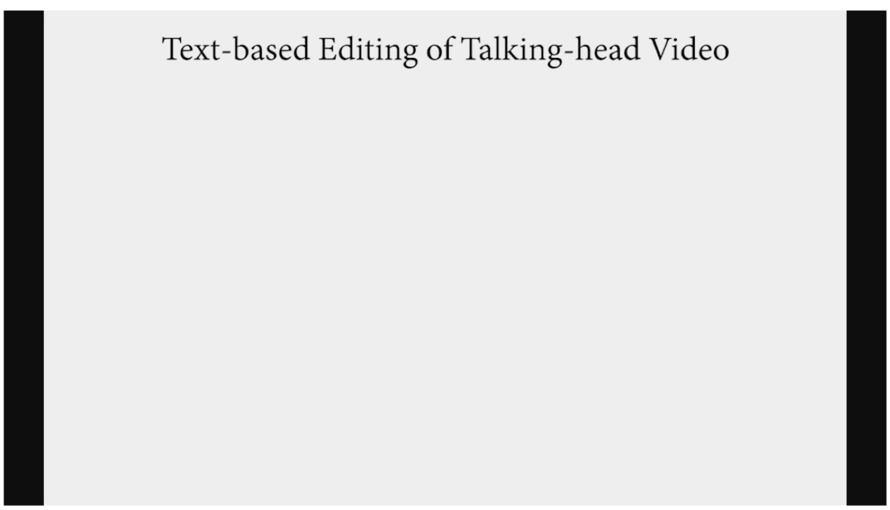
Course Goals

1. ***Gain overview of*** **algorithmic techniques used to manipulate video**
2. ***Present research paper*** **and** ***lead discussion*** **on a research paper**
3. ***Capture and edit*** **video manually and using algorithmic techniques**
4. ***Develop*** **substantial video manipulation project**

Instructor: Maneesh Agrawala

Visual Rhythm and Beat. Abe Davis and Maneesh Agrawala, SIGGRAPH 2018.

Instructor: Ohad Fried



Text-Based Editing of Talking Head Video. Ohad Fried, Ayush Tewari, Michael Zollhoefer, Adam Finkelstein, Eli Shectman, Dan B Goldman, Kyle Genova, Zeyu Jin, Christian Theobolt and Maneesh Agrawala, SIGGRAPH 2019.

Instructor: Michael Zollhöfer

Deep Video Portraits H. Kim, P. Garrido, A. Tewari, W. Xu, J. Thies, M. Nießner, P. Perez, C. Richardt, M. Zollhöfer, C. Theobalt SIGGRAPH 2018

Course Mechanics

Readings, Discussions, Presentations

Required to read about one paper per class

We will provide prompts to guide reading

You are responsible for written response to prompt

Due on paper at beginning of class, 2 free passes for the quarter

Required to present a paper and lead discussion once in the quarter

Usually Mon will be student presentations

You will meet with us (instructors) in week before presentation to go over 1st draft

Website

https://magrawala.github.io/cs448v-sp19/

COMPUTATIONAL VIDEO MANIPULATION

Stanford course CS448V focusing on algorithmic techniques for manipulating video (Spr 2019)

Location: Gates 176
Time: MW 1:30-2:50pm

SCHEDULE

Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10

TEACHING STAFF

ASSIGNMENTS

Visual media in the form of television programs, online video, and cinematic films have the power to engage people with dynamic, presentations of ideas. Expert storytellers design how such media unfolds over time to help audiences make sense of complex concepts, appreciate cultural or societal differences and imagine living in entirely different worlds. Technological advances have made it cheaper and easier to capture audio-visual media using the video cameras that are readily available in our mobile and desktop devices. Yet, the most viewed video are not simply raw recordings thrown onto the Web. The best material is carefully composed, filtered and edited to ensure that the resulting media is clear and engaging.

Nevertheless, today's tools for authoring and viewing video treat the media as a "baked" stream of audio samples, pixels, and frames – the very lowest-level representation possible. They have no understanding of the higher-level semantic structure of the audio-visual content. Researchers have developed a variety of techniques for extracting such higher-level structure from video and shown how to use this structure to significantly facilitate analysis, browsing, editing and manipulation of the raw material.

The goal of this graduate seminar (advanced undergraduates also welcome) is to survey recent work on computational video analysis and manipulation techniques. We will learn how to acquire, represent, edit and remix video. Several popular video manipulation algorithms will be presented, with an emphasis on using these techniques to build practical systems. Students will have the opportunity to acquire their own video and develop the processing tools needed to computationally analyze and manipulate it.

There are no official prerequisites for the course, but we will expect familiarity with the basic concepts of Computer Graphics and/or Computer Vision at the level of CS 148/248 and/or CS 131. Contact me (Maneesh) via email if you are worried about whether you have the background for the course.

Schedule

Week 1

M Apr 1: Introduction/Feature Detection and Tracking

Optional readings

Chapter 4.1: Feature Detection and Matching: Points and Patches. Szeliski. 2010. (pdf)

W Apr 3: Warping/RANSAC/Morphing

Required readings

Feature-Based Image Metamorphosis. Beier and Neely. SIGGRAPH 1992. (pdf)

Michael Jackson's Black or White video, morphing sequence. (YouTube)

Optional readings

Chapter 2.1: Image Formation: Geometric primitives and transformations. Szeliski. 2010. (pdf)

Chapter 6.1: Feature-Based Alignment: 2D and 3D Feature-Based Alignment. Szeliski. 2010. (pdf)

Week 2

M Apr 8: Video Texture and Looping

Required readings

Video Textures. Schodl et al. SIGGRAPH 2000. (pdf)

Optional readings

Panoramic Video Textures. Agarwala et al. SIGGRAPH 2005. (pdf)

Requirements

Participation (15%)

Attendance with prompt response is mandatory (but 2 free passes)
Also must engage in discussion in class

Presentation (15%)

Deeply engage with at least one paper and help others understand it

Assignments (20%)

Will help you learn about manual editing and the programmatic toolkits (e.g. OpenCV) available to implement algorithms

Final Project (50%)

Implement a research project on video manipulation

A1: Manual Manipulation

Interview a classmate and capture on video for at least 15 minutes

Plan the interview questions ahead of time
Capture on video (at least 15 minutes) – Do **not** hold camera, use a stand

Edit raw footage into a short video (< 2min) you would be proud to share

Use any video editing software you wish (e.g. Premiere, FinalCut Pro, iMovie)

Write down your reflections (half page PDF)

What was difficult in capturing and especially editing?
List all the pain points.
Describe ways video editing could be improved

Due Wed Apr 10 at 1:30pm

Feature Detection

Image Matching

by Diva Sian

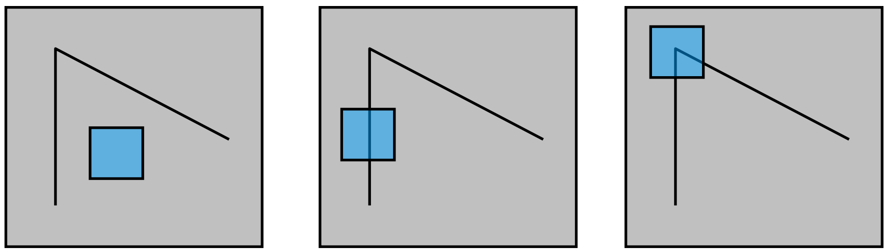
by scgbt

Slide credit: Seitz

Local Measures of Distinctiveness

Suppose we only consider a small window of pixels

What defines whether a feature is a good or bad candidate?

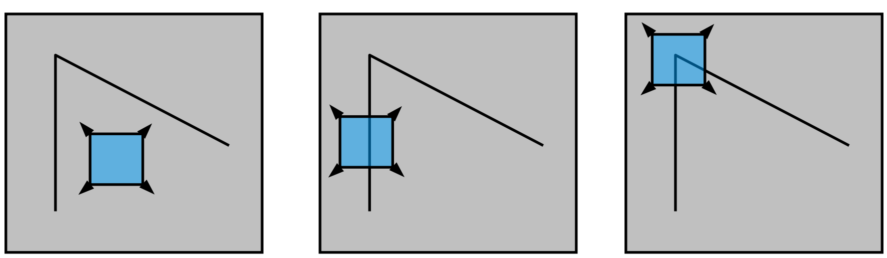


Slide credit: Seitz, Frovola, Simakov

Feature Detection

Local measure of feature uniqueness

- How does the window change when you shift it?
- Shifting the window in *any direction* causes a *big change*



"flat" region:
no change in all directions

"edge":
no change along the edge direction

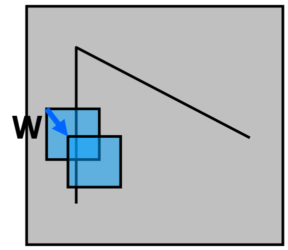
"corner":
significant change in all directions

Slide credit: Seitz, Frovola, Simakov

Feature Detection: Math

Consider shifting the window W by (u,v)

- How do the pixels in **W** change?
- Compare each pixel before and after by summing up the squared differences (SSD)
- This defines an SSD "error" of *E(u,v)*:



$$E(u,v) = \sum_{(x,y)\in W} \left[I(x+u,y+v) - I(x,y)\right]^2$$

Slide credit: Seitz, Frovola, Simakov

Small Motion Assumption

Taylor Series expansion of I:

$$I(x+u,y+v) = I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \text{higher order terms}$$

If the motion (u,v) is small, then first order approx is good

$$I(x+u,y+v) \approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v$$

$$\approx I(x,y) + [I_x \;\; I_y] \begin{bmatrix} u \\ v \end{bmatrix}$$

shorthand: $I_x = \frac{\partial I}{\partial x}$

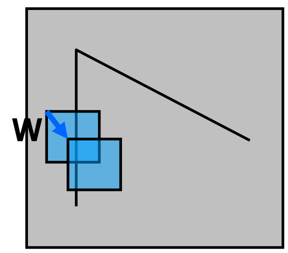
Plugging this into the formula on the previous slide…

Slide credit: Seitz, Frovola, Simakov

Feature Detection: Math

Consider shifting the window W by (u,v)

- How do the pixels in **W** change?
- Compare each pixel before and after by summing up the squared differences (SSD)
- This defines an SSD "error" of *E(u,v)*:



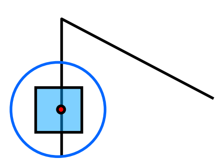
$$
\begin{aligned}
E(u,v) &= \sum_{(x,y)\in W} \left[I(x+u,y+v) - I(x,y)\right]^2 \\
&\approx \sum_{(x,y)\in W} \left[I(x,y) + [I_x \ I_y]\begin{bmatrix} u \\ v \end{bmatrix} - I(x,y)\right]^2 \\
&\approx \sum_{(x,y)\in W} \left[[I_x \ I_y]\begin{bmatrix} u \\ v \end{bmatrix}\right]^2
\end{aligned}
$$

Slide credit: Seitz, Frovola, Simakov

Feature Detection: Math

This can be rewritten:

$$E(u,v) = \sum_{(x,y)\in W} [u \ v] \underbrace{\begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$



Suppose you can move the center of the blue window in any direction

- Which directions will result in the largest and smallest E values?
- We can find these directions by looking at the eigenvectors of ***H***

Slide credit: Seitz, Frovola, Simakov

Eigenvalues & Eigenvectors

The **eigenvectors** of a matrix **A** are the vectors **x** that satisfy:

$$Ax = \lambda x$$

The scalar λ is the **eigenvalue** corresponding to **x**

The eigenvalues are found by solving:

$$det(A - \lambda I) = 0$$

- In our case, $\boldsymbol{A} = \boldsymbol{H}$ is a 2x2 matrix, so we have

$$det \begin{bmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{bmatrix} = 0$$

- The solution:

$$\lambda_{\pm} = \frac{1}{2}\left[(h_{11} + h_{22}) \pm \sqrt{4h_{12}h_{21} + (h_{11} - h_{22})^2}\right]$$

Once you know λ, you find **x** by solving

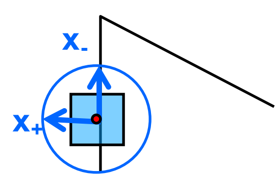
$$\begin{bmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

Slide credit: Seitz, Frovola, Simakov

Feature Detection: Math

This can be rewritten:

$$E(u,v) = \sum_{(x,y)\in W} [u\ v] \underbrace{\begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$



Eigenvalues and eigenvectors of H

- Define shifts with the smallest and largest change (E value)
- x_+ = direction of **largest** increase in E.
- λ_+ = amount of increase in direction x_+
- x_- = direction of **smallest** increase in E.
- λ_- = amount of increase in direction x_+

$$Hx_+ = \lambda_+ x_+$$
$$Hx_- = \lambda_- x_-$$

Slide credit: Seitz, Frovola, Simakov

Feature Detection: Math

How are λ_+, $\mathbf{x}_+$, λ_-, and $\mathbf{x}_+$ relevant for feature detection?

- What's our feature scoring function?

Slide credit: Seitz, Frovola, Simakov

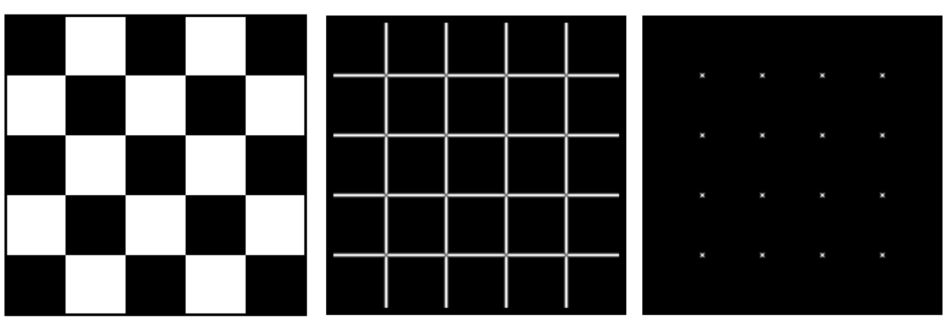
Feature detection: the math

How are λ_+, $\mathbf{x}_+$, λ_-, and $\mathbf{x}_+$ relevant for feature detection?

- What's our feature scoring function?

Want *E(u,v)* to be ***large*** for small shifts in ***all*** directions

- the *minimum* of *E(u,v)* should be large, over all unit vectors [u v]
- this minimum is given by the smaller eigenvalue (λ_-) of ***H***



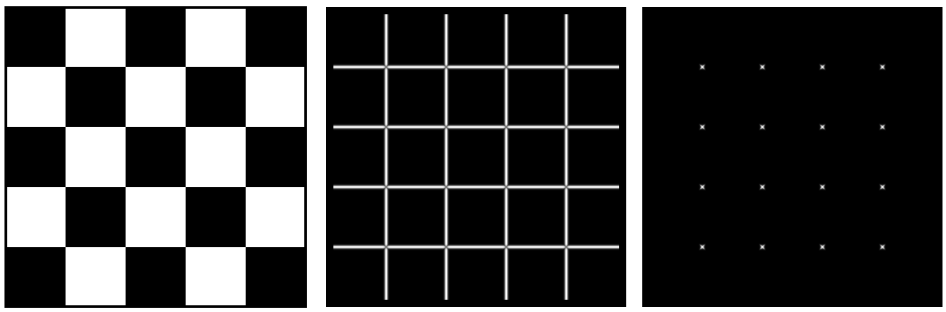
I λ_+ λ_-

Slide credit: Seitz, Frovola, Simakov

Feature Detection Summary

Here's what you do

- Compute the gradient at each point in the image
- Create the ***H*** matrix from the entries in the gradient
- Compute the eigenvalues.
- Find points with large response (λ_- > threshold)
- Choose those points where λ_- is a local maximum as features



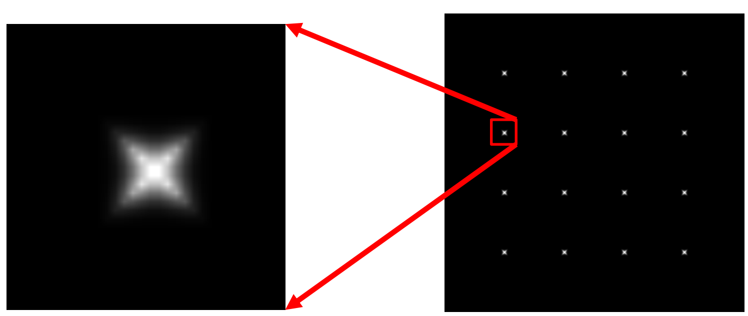
I λ_+ λ_-

Slide credit: Seitz, Frovola, Simakov

Feature Detection Summary

Here's what you do

- Compute the gradient at each point in the image
- Create the ***H*** matrix from the entries in the gradient
- Compute the eigenvalues.
- Find points with large response (λ_- > threshold)
- Choose those points where λ_- is a local maximum as features



λ_-

Slide credit: Seitz, Frovola, Simakov

The Harris Operator

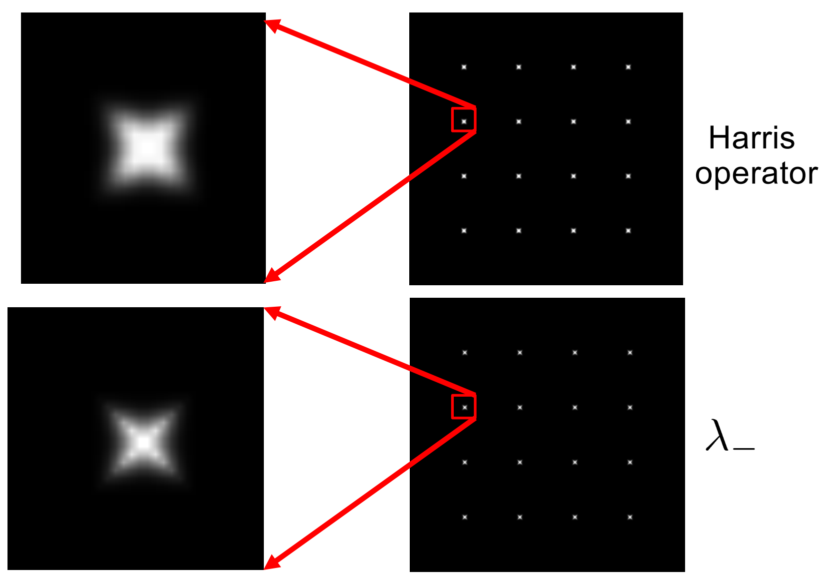
λ_- is a variant of the "Harris operator" for feature detection

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$
$$= \frac{determinant(H)}{trace(H)}$$

- The *trace* is the sum of the diagonals, i.e., $trace(H) = h_{11} + h_{22}$
- Very similar to λ_- but less expensive (no square root)
- Called the "Harris Corner Detector" or "Harris Operator"
- Lots of other detectors, this is one of the most popular

Slide credit: Seitz, Frovola, Simakov

The Harris Operator

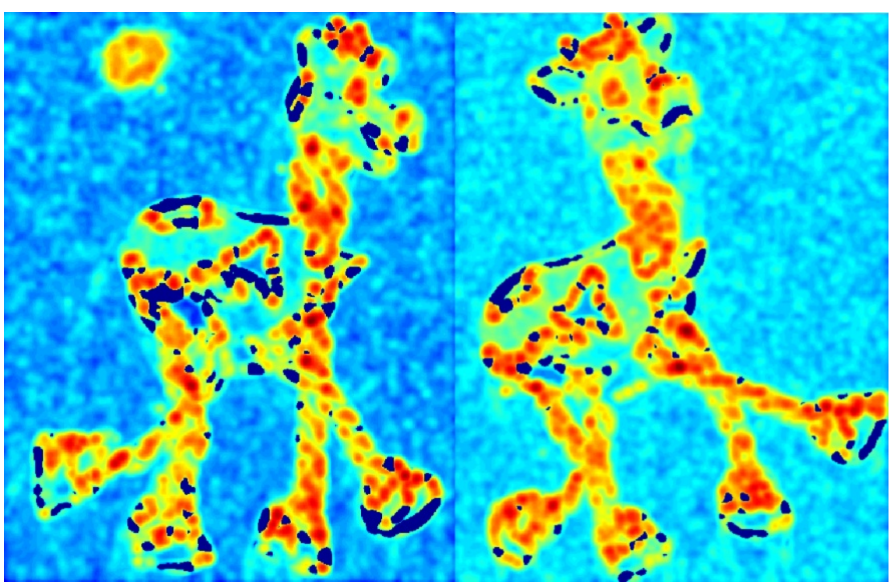


Slide credit: Seitz, Frovola, Simakov

Harris Operator Example

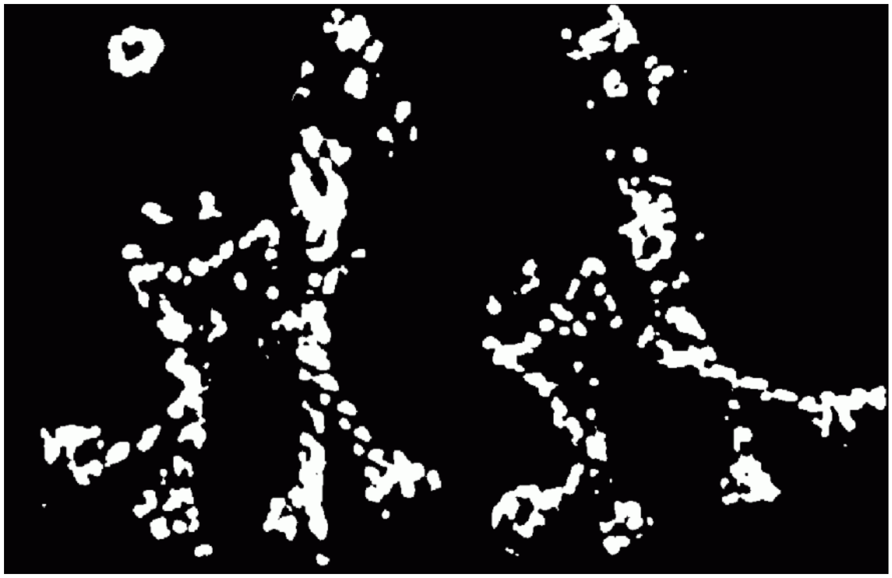
Slide credit: Seitz, Frovola, Simakov

f value (red high, blue low)



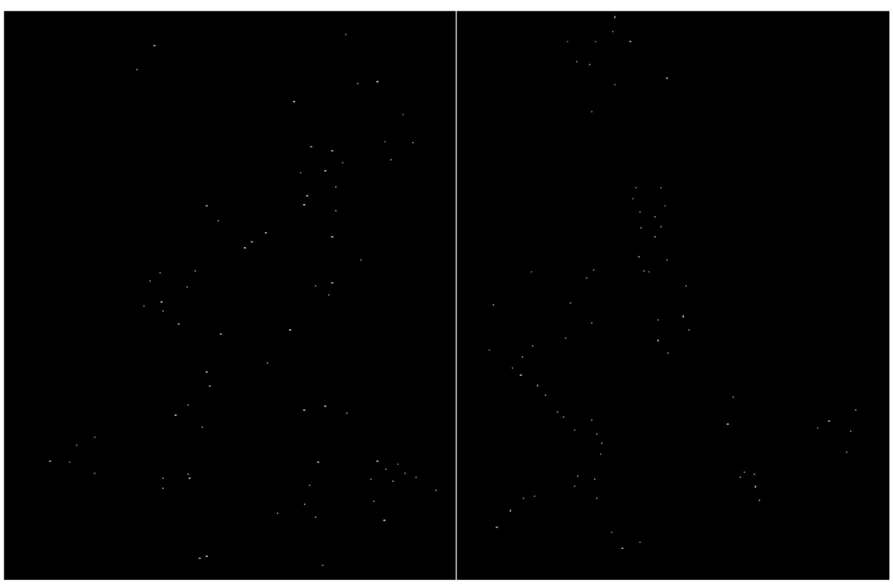
Slide credit: Seitz, Frovola, Simakov

Threshold (f > value)



Slide credit: Seitz, Frovola, Simakov

Find Local Maxima of f



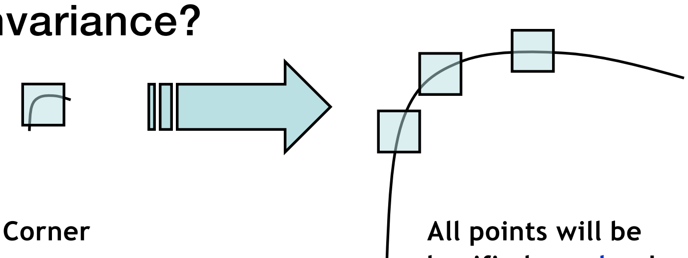
Slide credit: Seitz, Frovola, Simakov

Harris Features (in red)

Slide credit: Seitz, Frovola, Simakov

Invariance with Harris Corners

- Translation invariance
- Rotation invariance
- Scale invariance?



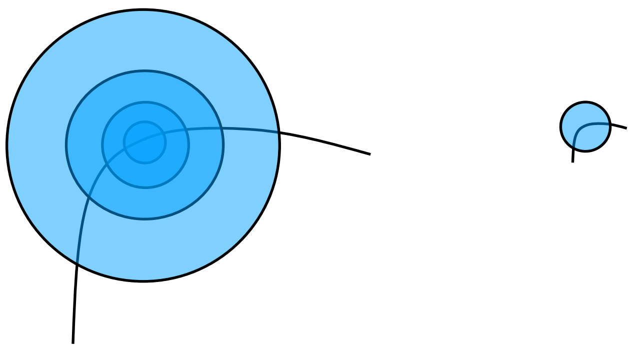
Not invariant to image scale!

Slide credit: Kristen Grauman

Scale Invariant Detection

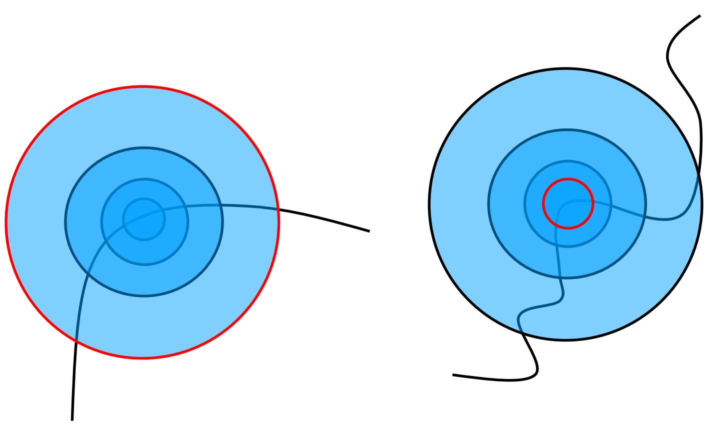
Consider regions (e.g. circles) of different sizes around a point

Find regions of corresponding sizes that will look the same in both images?

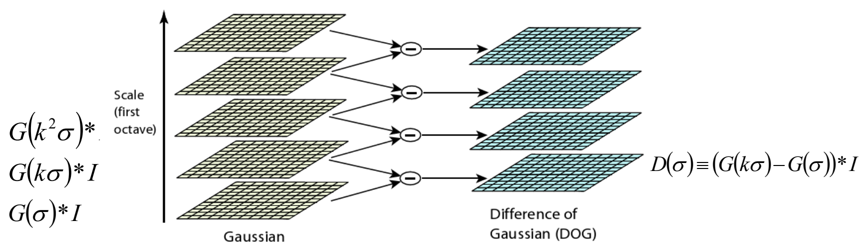


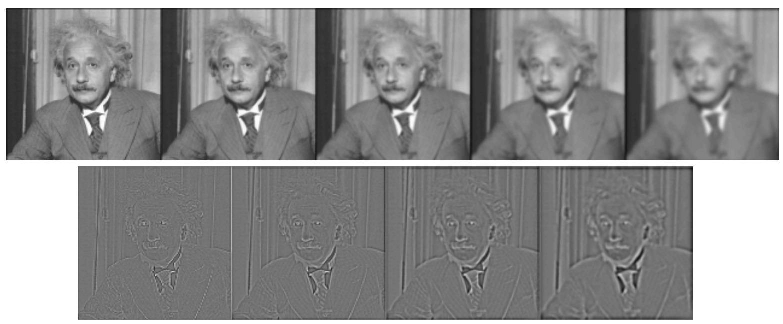
Scale Invariant Detection

The problem: how do we choose corresponding circles ***independently*** in each image?



Difference of Gaussians

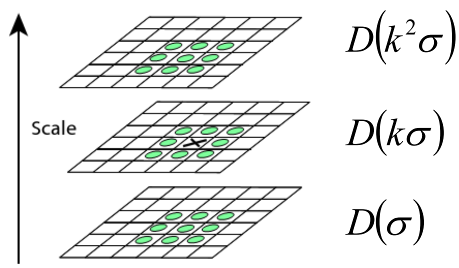




Slide credit: Niebles and Krishna

Scale-Space Extrema

Choose all extrema within 3x3x3 neighborhood



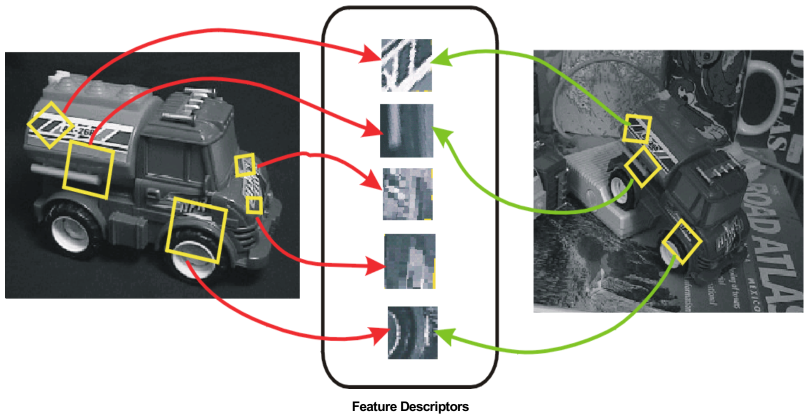
X is selected if it is larger or smaller than all 26 neighbors

Slide credit: Niebles and Krishna

Invariant Local Features

Find features that are invariant to transformations

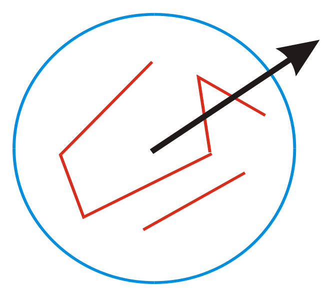
- geometric invariance: translation, rotation, scale
- photometric invariance: brightness, exposure, …



Slide credit: Niebles and Krishna

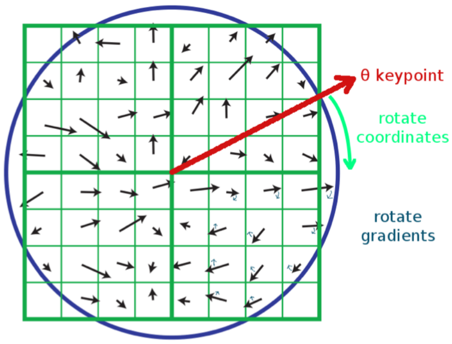
Becoming Rotation Invariant

- We are given a keypoint and its scale from DoG
- We select a characteristic orientation for the keypoint (based on the most prominent gradient in local region)
- We describe all features **relative** to this orientation
- Causes features to be rotation invariant!
  - If the keypoint appears rotated in another image, the features will be the same, because they're **relative** to the characteristic orientation



Slide credit: Niebles and Krishna

SIFT Descriptor Formation



Use the blurred image associated with the keypoint's scale

Take image gradients over the keypoint 16x16 neighborhood (put in 36 bin histogram)

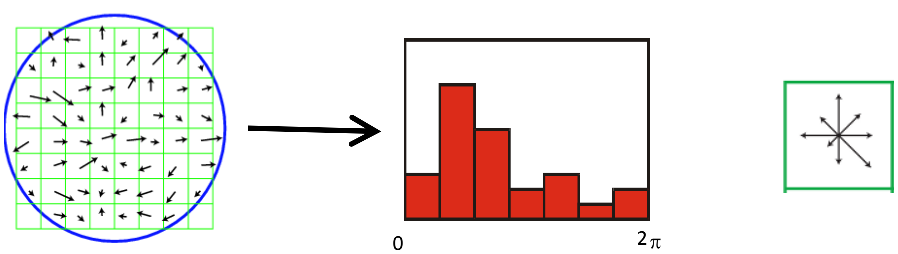
- Treat max bin as keypoint orientation θ

To become rotation invariant, rotate the gradient directions AND locations by (- θ)

- Now we've cancelled out rotation and have gradients expressed at locations **relative** to keypoint orientation θ
- We could also have just rotated the whole image by -θ, but that would be slower

Slide credit: Niebles and Krishna

SIFT Descriptor Formation



Using precise gradients & locations is fragile

For robustness create array of orientation histograms

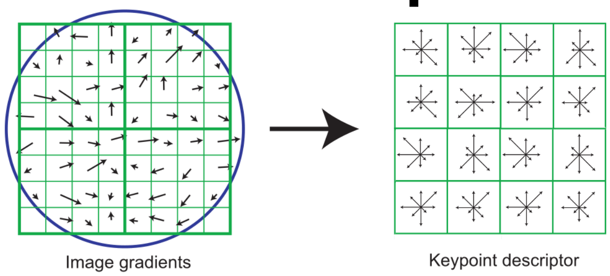
Put the rotated gradients into their local orientation histograms

- A gradient's contribution is divided among the nearby histograms based on distance. If it's halfway between two histogram locations, it gives a half contribution to both.
- Also, scale down gradient contributions for gradients far from the center

The SIFT authors found that best results were with 8 orientation bins per histogram

Slide credit: Niebles and Krishna

SIFT Descriptor Formation



Using precise gradients & locations is fragile

For robustness create array of orientation histograms

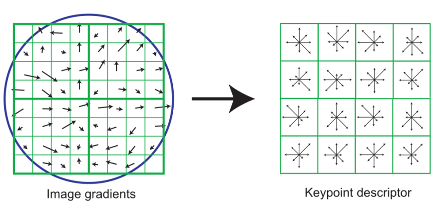
Put the rotated gradients into their local orientation histograms

- A gradient's contribution is divided among the nearby histograms based on distance. If it's halfway between two histogram locations, it gives a half contribution to both.
- Also, scale down gradient contributions for gradients far from the center

The SIFT authors found that best results were with 8 orientation bins per histogram and 4x4 histogram array

Slide credit: Niebles and Krishna

SIFT Descriptor Formation



8 orientation bins per histogram, and 4x4 histogram array: 8 x 4x4 = 128 numbers

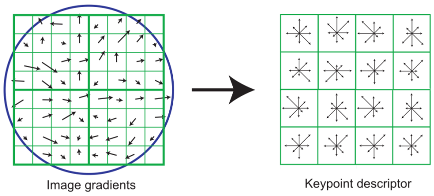
So a SIFT descriptor is a length 128 vector, which is invariant to rotation (because we rotated the descriptor) and scale (because we worked with the scaled image from DoG)

We can compare each vector from image A to each vector from image B to find matching keypoints!

Euclidean "distance" between descriptor vectors gives a good measure of keypoint similarity

Slide credit: Niebles and Krishna

SIFT Descriptor Formation



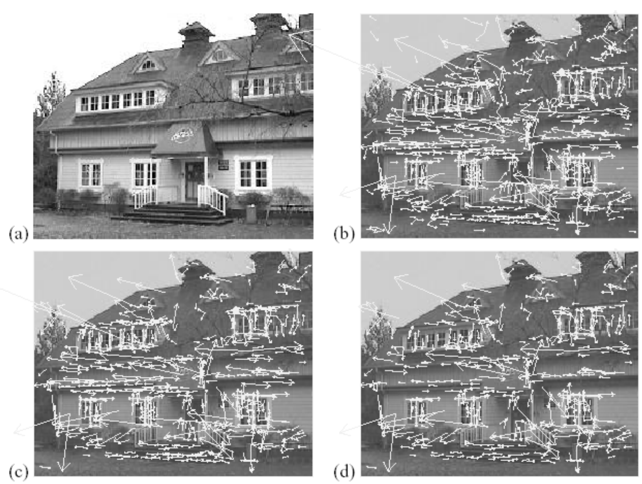
Adding robustness to illumination changes:

- Descriptor is made of gradients (differences between pixels), so it's already invariant to changes in brightness (e.g. adding 10 to all image pixels yields the exact same descriptor)
- A higher-contrast photo will increase the magnitude of gradients linearly. So, to correct for contrast changes, normalize the vector (scale to length 1.0)
- Very large image gradients are usually from unreliable 3D illumination effects (glare, etc). So, to reduce their effect, clamp all values in the vector to be ≤ 0.2 (an experimentally tuned value). Then normalize the vector again.

Slide credit: Niebles and Krishna

SIFT Keypoints Detection

Threshold on value at DOG peak and on ratio of principle curvatures



(a) 233x189 image
(b) 832 DOG extrema
(c) 729 left after peak value threshold
(d) 536 left after testing ratio of principle curvatures

Vectors indicate scale, orientation and location

Slide credit: Niebles and Krishna

Properties of SIFT

Extraordinarily robust matching technique

- Can handle changes in viewpoint (Up to about 60 degree out of plane rotation)
- Can handle significant changes in illumination (Sometimes even day vs. night (see below))
- Fast and efficient—can run in real time
- Lots of code available
  - http://people.csail.mit.edu/albert/ladypack/wiki/index.php/Known_implementations_of_SIFT

Slide credit: Seitz

Feature Matching

Given a feature in I_1, how to find the best match in I_2?

1. Define distance function that compares two descriptors
   (e.g. Euclidean distance between SIFT descriptors)

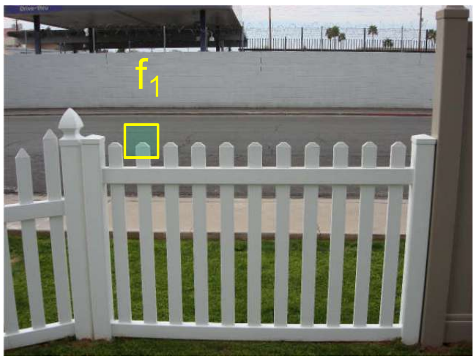
2. Test all the features in I_2, find the one with min distance

Slide credit: Seitz

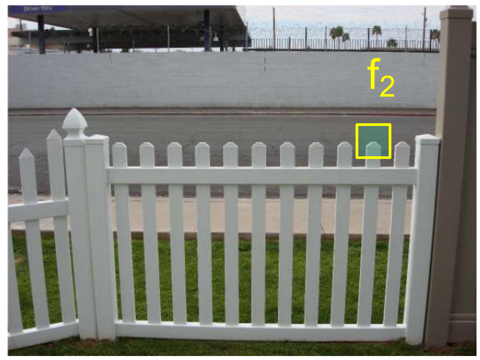
Feature Distance

How to define the difference between two features f_1, f_2?

- Simple approach is SSD(f_1, f_2)
  - sum of square differences between entries of the two descriptors
  - can give good scores to very ambiguous (bad) matches



I_1



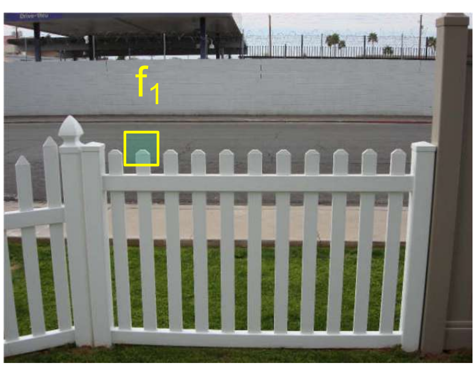
I_2

Slide credit: Seitz

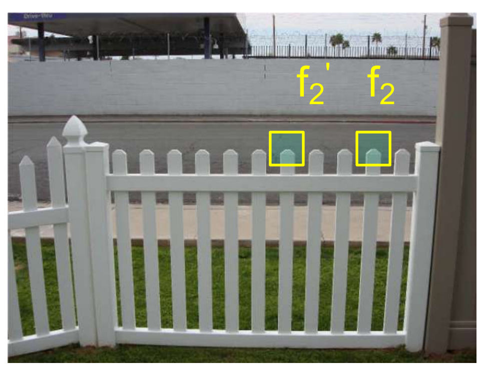
Feature Distance

How to define the difference between two features f_1, f_2?

- Better approach: ratio distance = SSD(f_1, f_2) / SSD(f_1, f_2')
  - f_2 is best SSD match to f_1 in I_2
  - f_2' is 2^{nd} best SSD match to f_1 in I_2
  - gives small values for ambiguous matches



I_1

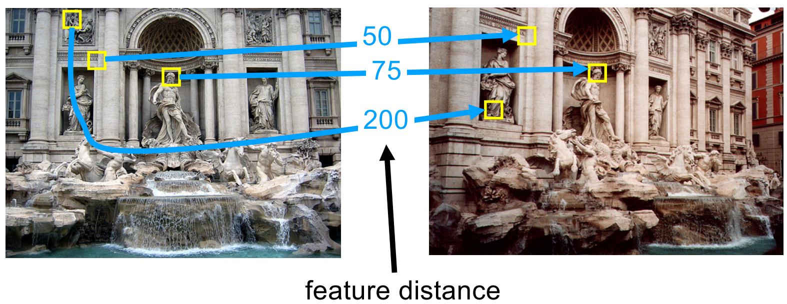


I_2

Slide credit: Seitz

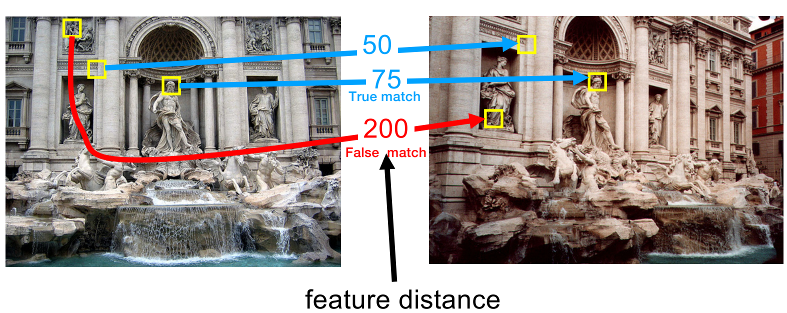
Evaluating the Results

How can we measure the performance of a feature matcher?



Slide credit: Seitz

True/False Positives

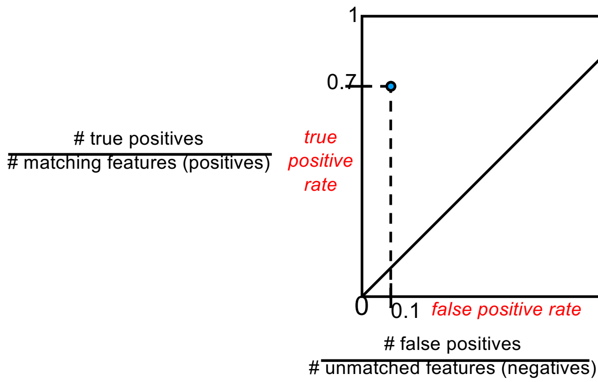


The distance threshold affects performance

- True positives = # of detected matches that are correct
  - Suppose we want to maximize these—how to choose threshold?
- False positives = # of detected matches that are incorrect
  - Suppose we want to minimize these—how to choose threshold?

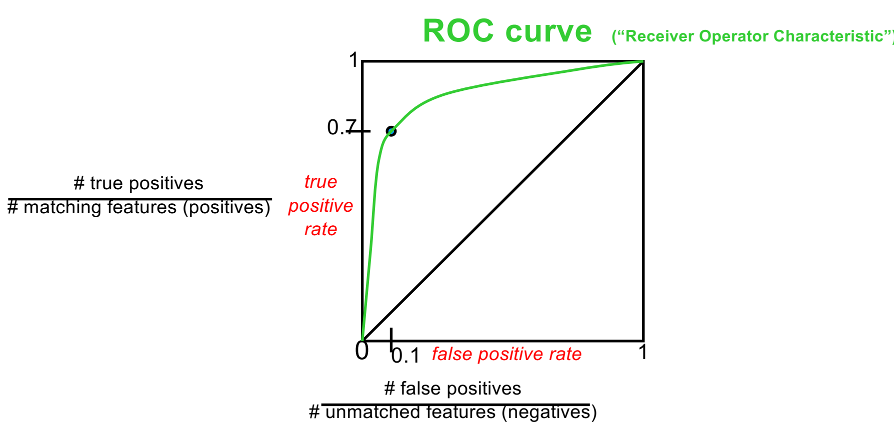
Slide credit: Seitz

Evaluating the Results



Slide credit: Seitz

Evaluating the Results

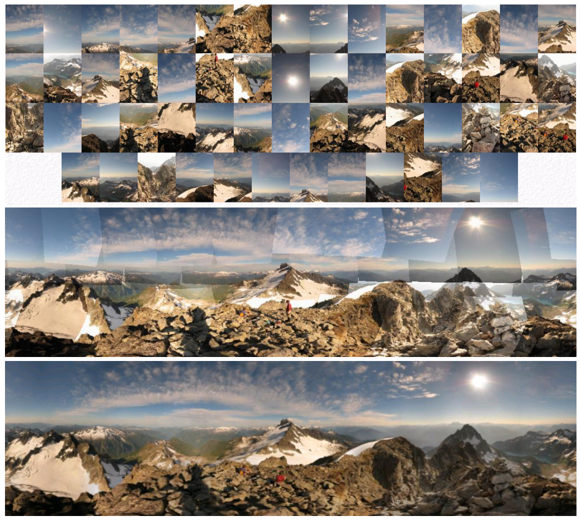


ROC Curves

- Generated by counting # current/incorrect matches, for different thresholds
- Want to maximize area under the curve (AUC)
- Useful for comparing different feature matching methods
- For more info: http://en.wikipedia.org/wiki/Receiver_operating_characteristic

Slide credit: Seitz

Application: Mosaicing



http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html

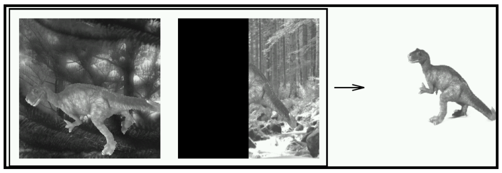
Slide credit: Niebles and Krishna

Application: Wide Baseline Stereo

[Image from T. Tuytelaars ECCV 2006 tutorial]

Slide credit: Niebles and Krishna

Application: Object/Scene Recognition


Schmid and Mohr 1997


Sivic and Zisserman, 2003


Rothganger et al. 2003


Lowe 2002

Slide credit: Niebles and Krishna