More on Features

Digital Visual Effects, Spring 2008

Yung-Yu Chuang

2008/3/25

with slides by Trevor Darrell Cordelia Schmid, David Lowe, Darya Frolova, Denis Simakov, Robert Collins, Brad Osgood, W W L Chen, and Jiwon Kim

Announcements

- Project #1 is due. Voting will start from next Tuesday.
- Project #2 handout is out. Due four weeks later. A checkpoint at 4/11.

Multi-Scale Oriented Patches

Multi-Scale Oriented Patches

- Simpler than SIFT. Designed specifically for image matching. [Brown, Szeliski and Winder, CVPR'2005]
- Feature detector
  - Multi-scale Harris corners
  - Orientation from blurred gradient
  - Geometrically invariant to rotation
- Feature descriptor
  - Bias/gain normalized sampling of local patch (8x8)
  - Photometrically invariant to affine changes in intensity

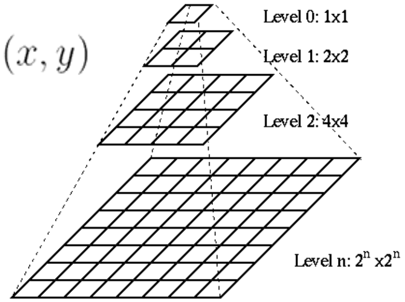
Multi-Scale Harris corner detector

$$P_0(x,y) = I(x,y)$$

$$P'_l(x,y) = P_l(x,y) * g_{\sigma_p}(x,y)$$

$$P_{l+1}(x,y) = P'_l(sx, sy)$$

$$s = 2 \qquad \sigma_p = 1.0$$



- Image stitching is mostly concerned with matching images that have the same scale, so sub-octave pyramid might not be necessary.

Multi-Scale Harris corner detector

$$\mathbf{H}_l(x,y) = \nabla_{\sigma_d} P_l(x,y) \nabla_{\sigma_d} P_l(x,y)^T * g_{\sigma_i}(x,y)$$

$$\nabla_\sigma f(x,y) \triangleq \nabla f(x,y) * g_\sigma(x,y)$$

smoother version of gradients

$$\sigma_i = 1.5 \qquad \sigma_d = 1.0$$

Corner detection function:

$$f_{HM}(x,y) = \frac{\det \mathbf{H}_l(x,y)}{\operatorname{tr} \mathbf{H}_l(x,y)} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

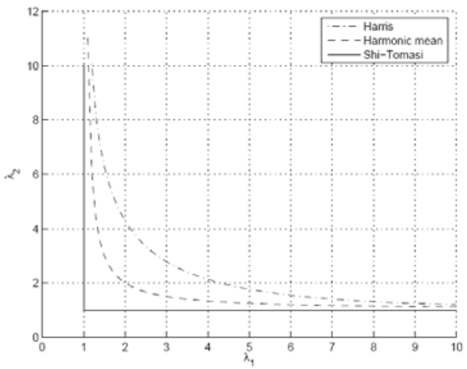
Pick local maxima of 3x3 and larger than 10

Keypoint detection function

$$\text{Harris} \quad f_H = \lambda_1\lambda_2 - 0.04(\lambda_1+\lambda_2)^2 = \det \mathbf{H} - 0.04(\operatorname{tr} \mathbf{H})^2$$

$$\text{Harmonic mean} \quad f_{HM} = \lambda_1\lambda_2/(\lambda_1+\lambda_2) = \det \mathbf{H}/\operatorname{tr} \mathbf{H}$$

$$\text{Shi-Tomasi} \quad f_{ST} = \min(\lambda_1, \lambda_2)$$

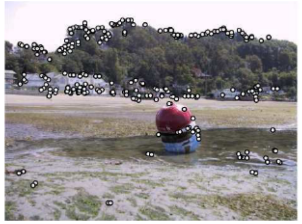


Experiments show roughly the same performance.

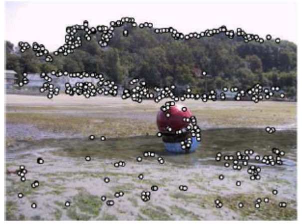
Non-maximal suppression

- Restrict the maximal number of interest points, but also want them spatially well distributed
- Only retain maximums in a neighborhood of radius r.
- Sort them by strength, decreasing r from infinity until the number of keypoints (500) is satisfied.

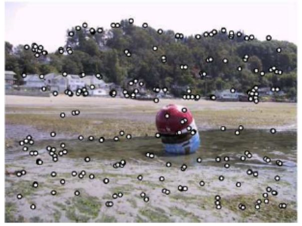
Non-maximal suppression



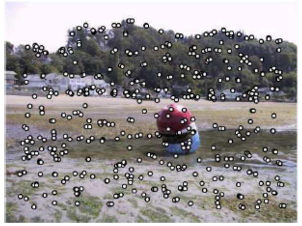
(a) Strongest 250



(b) Strongest 500



(c) ANMS 250, $r = 24$



(d) ANMS 500, $r = 16$

Sub-pixel refinement

$$f(\mathbf{x}) = f + \frac{\partial f}{\partial \mathbf{x}}^T \mathbf{x} + \frac{1}{2}\mathbf{x}^T \frac{\partial^2 f}{\partial \mathbf{x}^2}\mathbf{x}$$

$$\mathbf{x}_m = -\frac{\partial^2 f}{\partial \mathbf{x}^2}^{-1} \frac{\partial f}{\partial \mathbf{x}}$$

$f_{-1,1}$	$f_{0,1}$	$f_{1,1}$
$f_{-1,0}$	$f_{0,0}$	$f_{1,0}$
$f_{-1,-1}$	$f_{0,-1}$	$f_{1,-1}$

$$\begin{aligned}
\frac{\partial f}{\partial x} &= (f_{1,0} - f_{-1,0})/2 \\
\frac{\partial f}{\partial y} &= (f_{0,1} - f_{0,-1})/2 \\
\frac{\partial^2 f}{\partial x^2} &= f_{1,0} - 2f_{0,0} + f_{-1,0} \\
\frac{\partial^2 f}{\partial y^2} &= f_{0,1} - 2f_{0,0} + f_{0,-1} \\
\frac{\partial^2 f}{\partial x \partial y} &= (f_{-1,-1} - f_{-1,1} - f_{1,-1} + f_{1,1})/4
\end{aligned}$$

Orientation assignment

- Orientation = blurred gradient

$$\mathbf{u}_l(x, y) = \nabla_{\sigma_o} P_l(x, y)$$

$$\sigma_o = 4.5$$

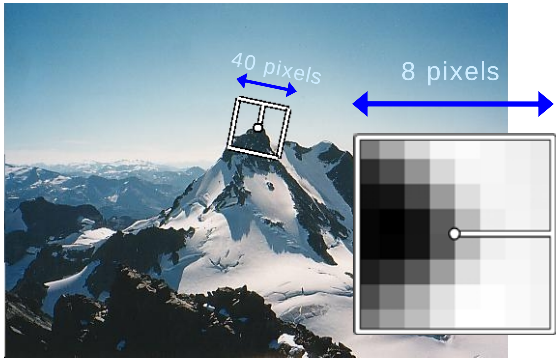
$$[\cos\theta, \sin\theta] = \mathbf{u}/|\mathbf{u}|$$

Descriptor Vector

- Rotation Invariant Frame
  - Scale-space position (x, y, s) + orientation (θ)

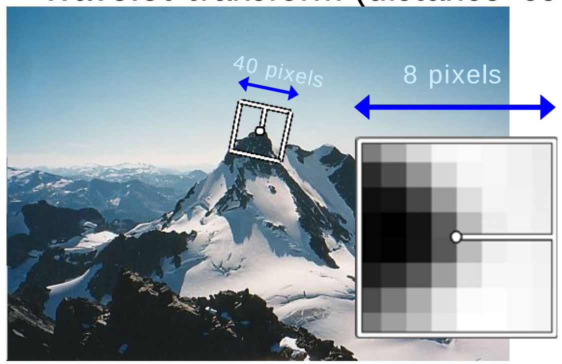
MOPS descriptor vector

- 8x8 oriented patch sampled at 5 x scale. See TR for details.
- Sampled from $P_l(x, y) * g_{2\times\sigma_p}(x, y)$ with spacing=5



MOPS descriptor vector

- 8x8 oriented patch sampled at 5 x scale. See TR for details.
- Bias/gain normalisation: $I' = (I - \mu)/\sigma$
- Wavelet transform (distance=coeff. distance)



Detections at multiple scales

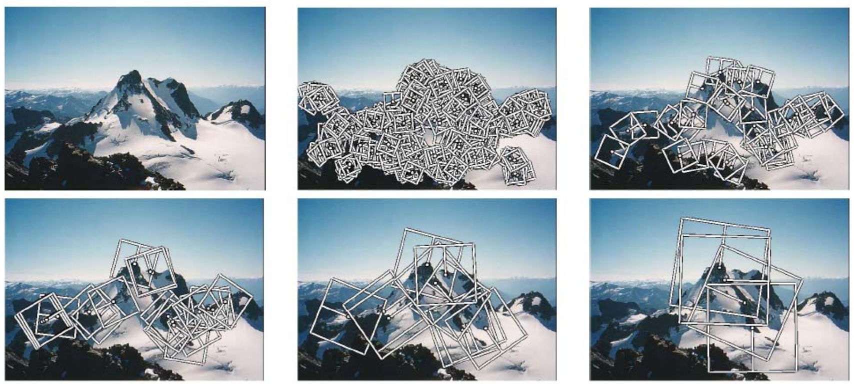


Figure 1. Multi-scale Oriented Patches (MOPS) extracted at five pyramid levels from one of the Matier images. The boxes show the feature orientation and the region from which the descriptor vector is sampled.

Summary

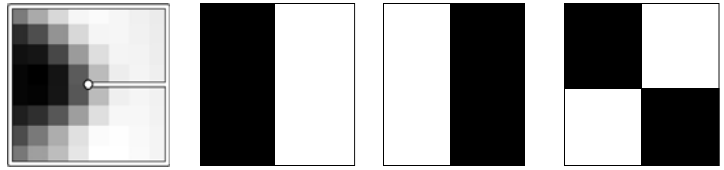
- Multi-scale Harris corner detector
- Sub-pixel refinement
- Orientation assignment by gradients
- Blurred intensity patch as descriptor

Feature matching

- Exhaustive search
  - for each feature in one image, look at *all* the other features in the other image(s)
- Hashing
  - compute a short descriptor from each feature vector, or hash longer descriptors (randomly)
- Nearest neighbor techniques
  - *k*-trees and their variants (Best Bin First)

Wavelet-based hashing

- Compute a short (3-vector) descriptor from an 8x8 patch using a Haar "wavelet"



- Quantize each value into 10 (overlapping) bins (10^3 total entries)
- [Brown, Szeliski, Winder, CVPR'2005]

Nearest neighbor techniques

- *k*-D tree and
- Best Bin First (BBF)

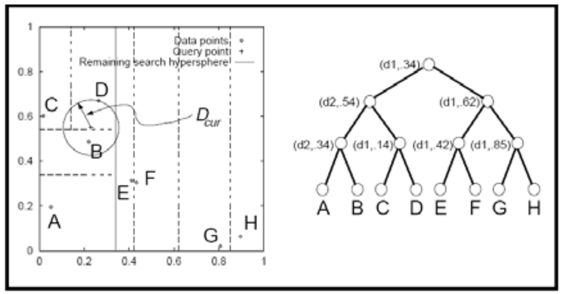
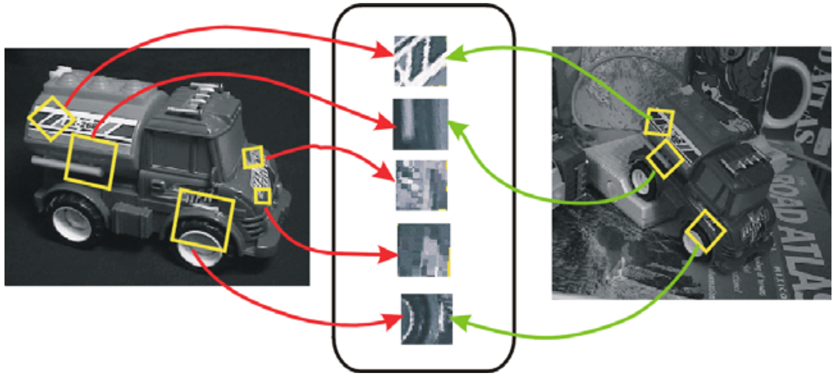


Figure 6: kd-tree with 8 data points labelled A-H, dimension of space k=2. On the right is the full tree, the leaf nodes containing the data points. Internal node information consists of the dimension of the cut plane and the value of the cut in that dimension. On the left is the 2D feature space carved into various sizes and shapes of bin, according to the distribution of the data points. The two representations are isomorphic. The situation shown on the left is after initial tree traversal to locate the bin for query point "+" (contains point D). In standard search, the closest nodes in the tree are examined first (starting at C). In BBF search, the closest bins to query point q are examined first (starting at B). The latter is more likely to maximize the overlap of (i) the hypersphere centered on q with radius D_{cur}, and (ii) the hyperrectangle of the bin to be searched. In this case, BBF search reduces the number of leaves to examine, since once point B is discovered, all other branches can be pruned.

Indexing Without Invariants in 3D Object Recognition, Beis and Lowe, PAMI'99

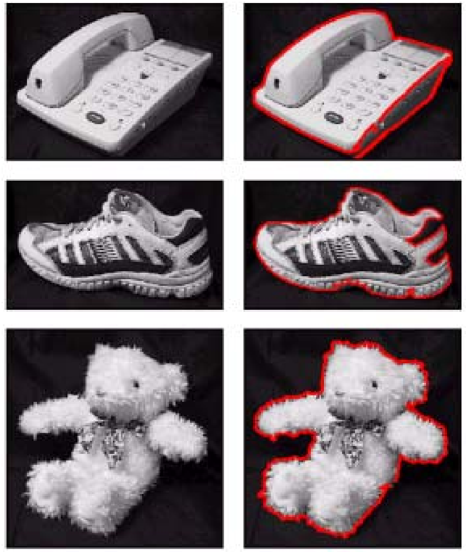
Applications of features

Recognition

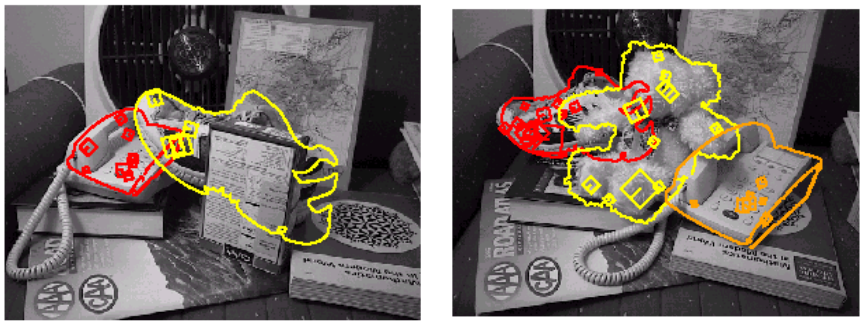


SIFT Features

3D object recognition



3D object recognition



Office of the past

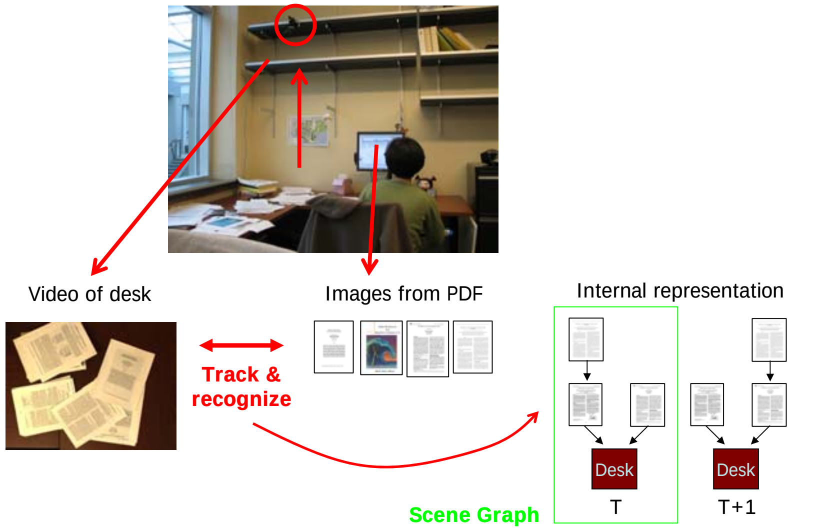
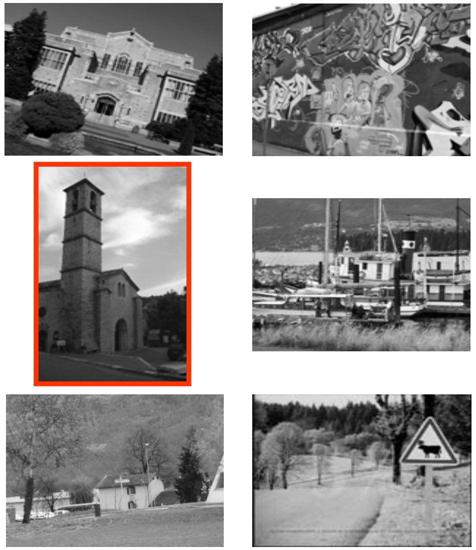


Image retrieval

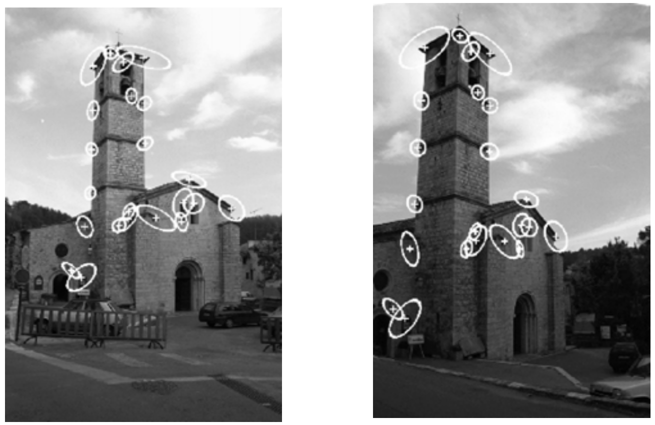
change in viewing angle



• • •

> 5000 images

Image retrieval



22 correct matches

Image retrieval

change in viewing angle
+ scale change

• • •

> 5000 images

Robot location

Robotics: Sony Aibo

SIFT is used for
- Recognizing charging station
- Communicating with visual cards
- Teaching object recognition
- soccer

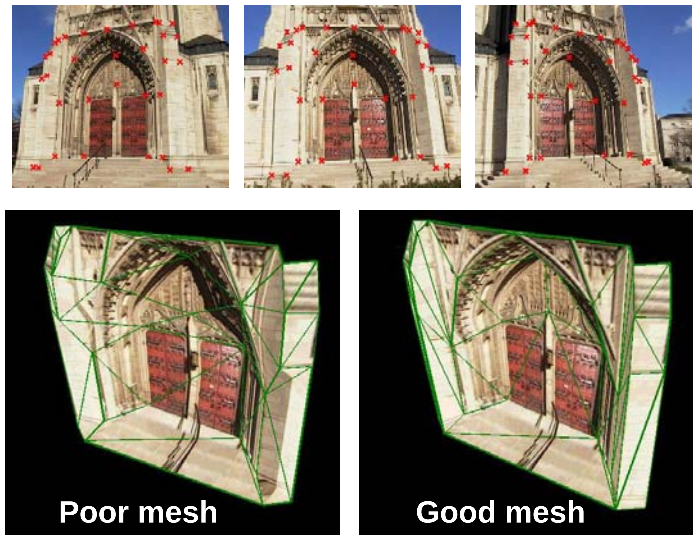
Structure from Motion

- The SFM Problem
  - Reconstruct scene geometry and camera motion from two or more images

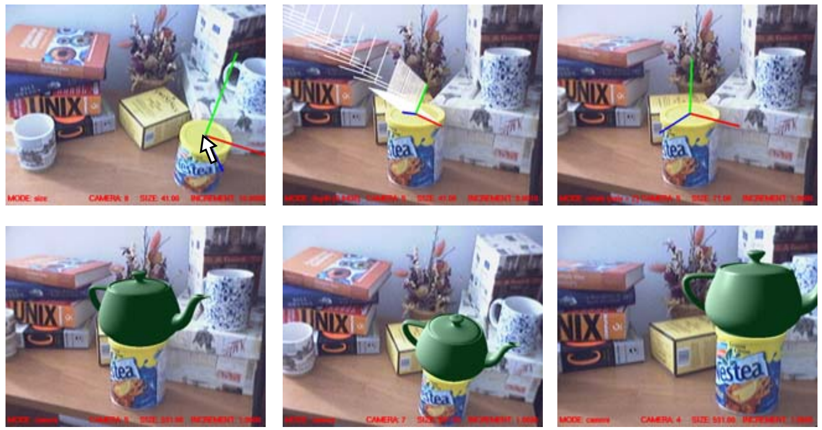
Track 2D Features → **Estimate 3D** → **Optimize (Bundle Adjust)** → **Fit Surfaces**

SFM Pipeline

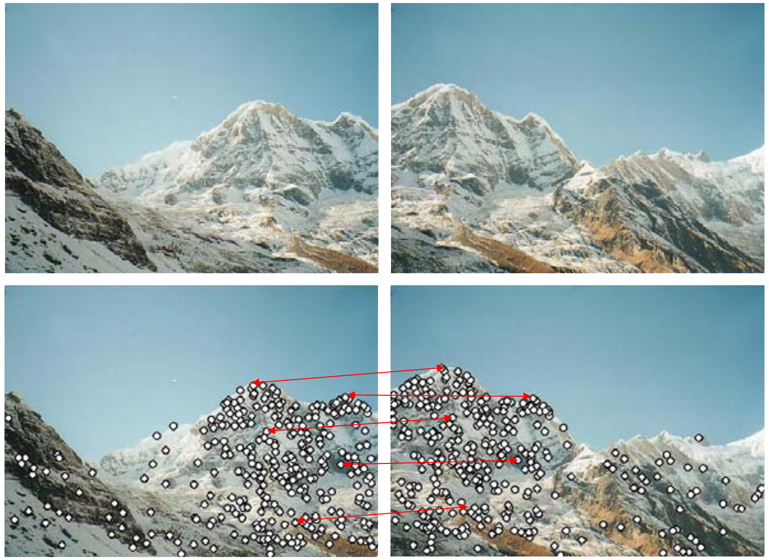
Structure from Motion



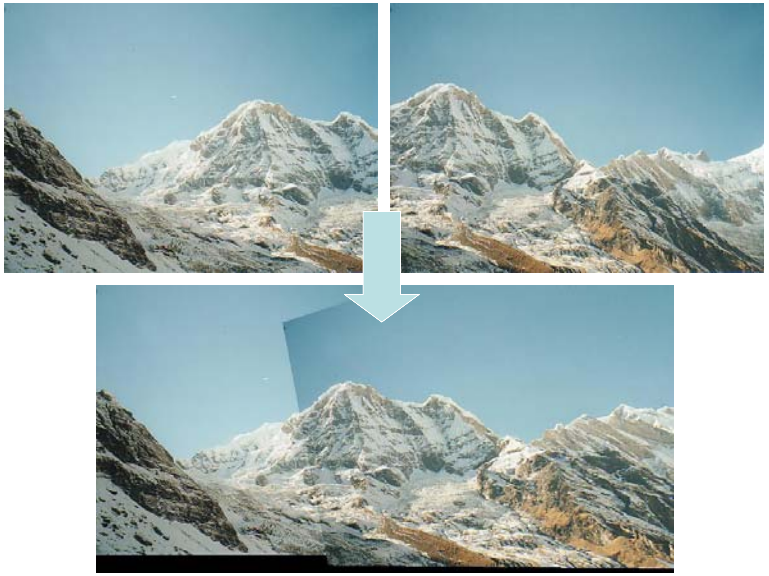
Augmented reality



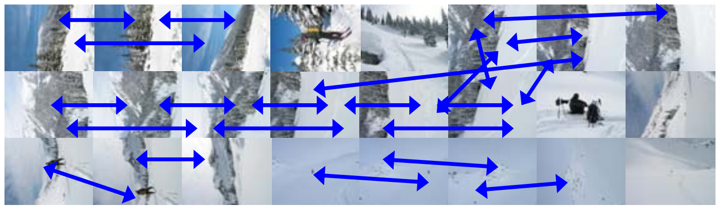
Automatic image stitching



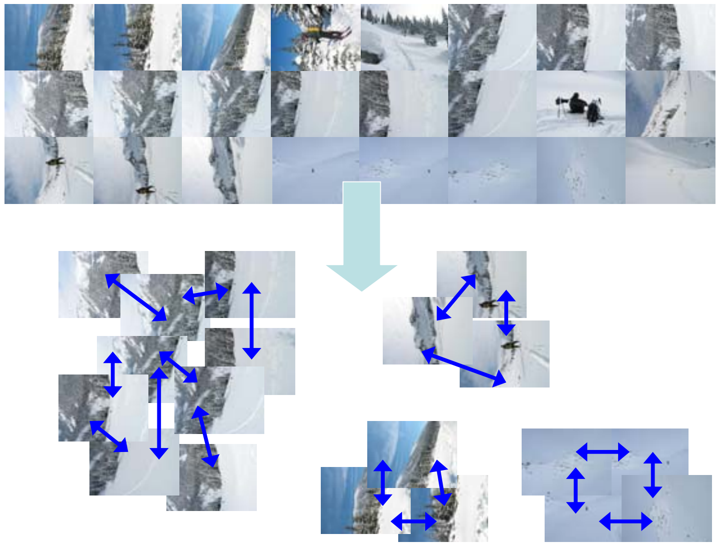
Automatic image stitching



Automatic image stitching



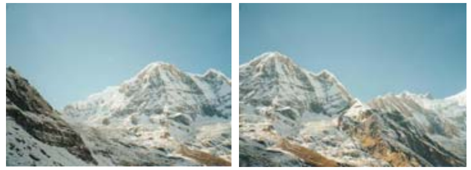
Automatic image stitching



Automatic image stitching

Project #2 Image stitching

- Assigned: 3/21
- Checkpoint: 11:59pm 4/11
- Due: 11:59pm 4/21
- Work in pairs



Reference software

- Autostitch

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

- Many others are available online.

Tips for taking pictures

- Common focal point
- Rotate your camera to increase vertical FOV
- Tripod
- Fixed exposure?

Bells & whistles

- Recognizing panorama
- Bundle adjustment
- Handle dynamic objects
- Better blending techniques

Artifacts

- Take your own pictures and generate a stitched image, be creative.
  - http://www.cs.washington.edu/education/courses/cse590ss/01wi/projects/project1/students/allen/index.html

Submission

- You have to turn in your complete source, the executable, a html report and an artifact.
- Report page contains:

  description of the project, what do you learn, algorithm, implementation details, results, bells and whistles…
- Artifacts must be made using your own program.

Reference

- Chris Harris, Mike Stephens, A Combined Corner and Edge Detector, 4th Alvey Vision Conference, 1988, pp147-151.
- David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, 60(2), 2004, pp91-110.
- Yan Ke, Rahul Sukthankar, PCA-SIFT: A More Distinctive Representation for Local Image Descriptors, CVPR 2004.
- Krystian Mikolajczyk, Cordelia Schmid, A performance evaluation of local descriptors, Submitted to PAMI, 2004.
- SIFT Keypoint Detector, David Lowe.
- Matlab SIFT Tutorial, University of Toronto.