Automatic Image Stitching using Invariant Features

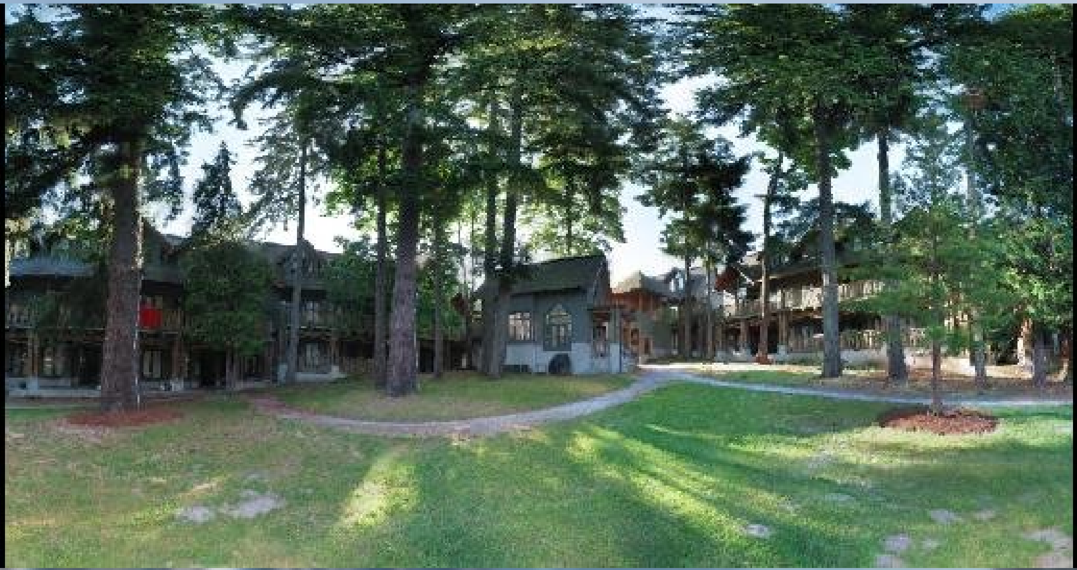
Matthew Brown and David Lowe,
University of British Columbia

Introduction

- Are you getting the whole picture?
  - Compact Camera FOV = 50 x 35°

Introduction

- Are you getting the whole picture?
  - Compact Camera FOV = 50 x 35°
  - Human FOV = 200 x 135°



Introduction

- Are you getting the whole picture?
  - Compact Camera FOV = 50 x 35°
  - Human FOV = 200 x 135°
  - Panoramic Mosaic = 360 x 180°

Recognising Panoramas

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

- 2D Rotations (θ, ϕ)
  - Ordering $\not\Rightarrow$ matching images

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

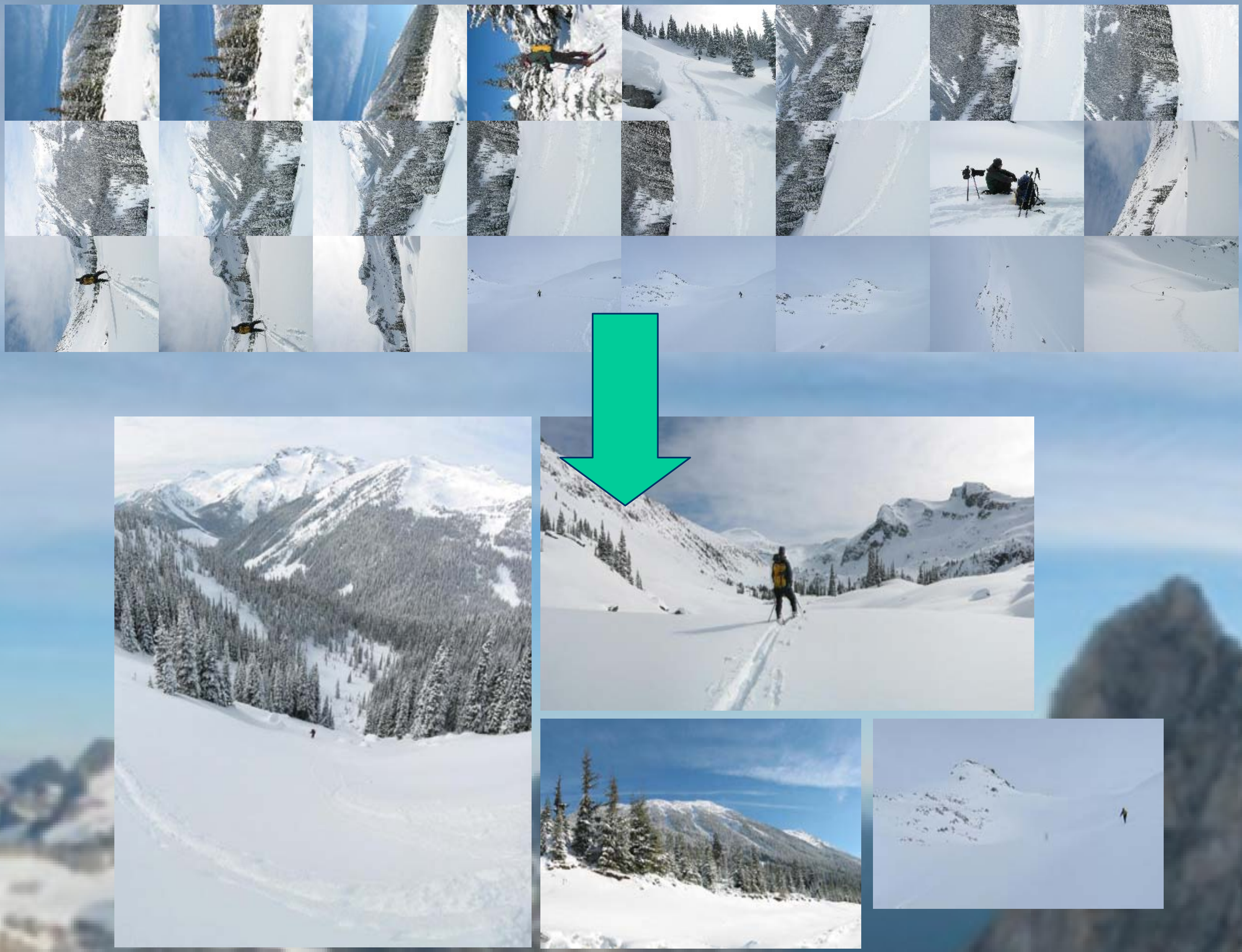
- 2D Rotations (θ, ϕ)
  - Ordering $\not\Rightarrow$ matching images

Recognising Panoramas

- 1D Rotations (θ)
  - Ordering $\Rightarrow$ matching images

- 2D Rotations (θ, ϕ)
  - Ordering $\not\Rightarrow$ matching images

Recognising Panoramas



Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
  - SIFT Features
  - Nearest Neighbour Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
  - SIFT Features
  - Nearest Neighbour Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Invariant Features

- Schmid & Mohr 1997, Lowe 1999, Baumberg 2000, Tuytelaars & Van Gool 2000, Mikolajczyk & Schmid 2001, Brown & Lowe 2002, Matas et. al. 2002, Schaffalitzky & Zisserman 2002

SIFT Features

- Invariant Features
  - Establish invariant frame
    - Maxima/minima of scale-space DOG $\Rightarrow$ x, y, s
    - Maximum of distribution of local gradients $\Rightarrow \theta$
  - Form descriptor vector
    - Histogram of smoothed local gradients
    - 128 dimensions
- SIFT features are…
  - Geometrically invariant to similarity transforms,
    - some robustness to affine change
  - Photometrically invariant to affine changes in intensity

Overview

- Feature Matching
  - SIFT Features
  - Nearest Neighbour Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Nearest Neighbour Matching

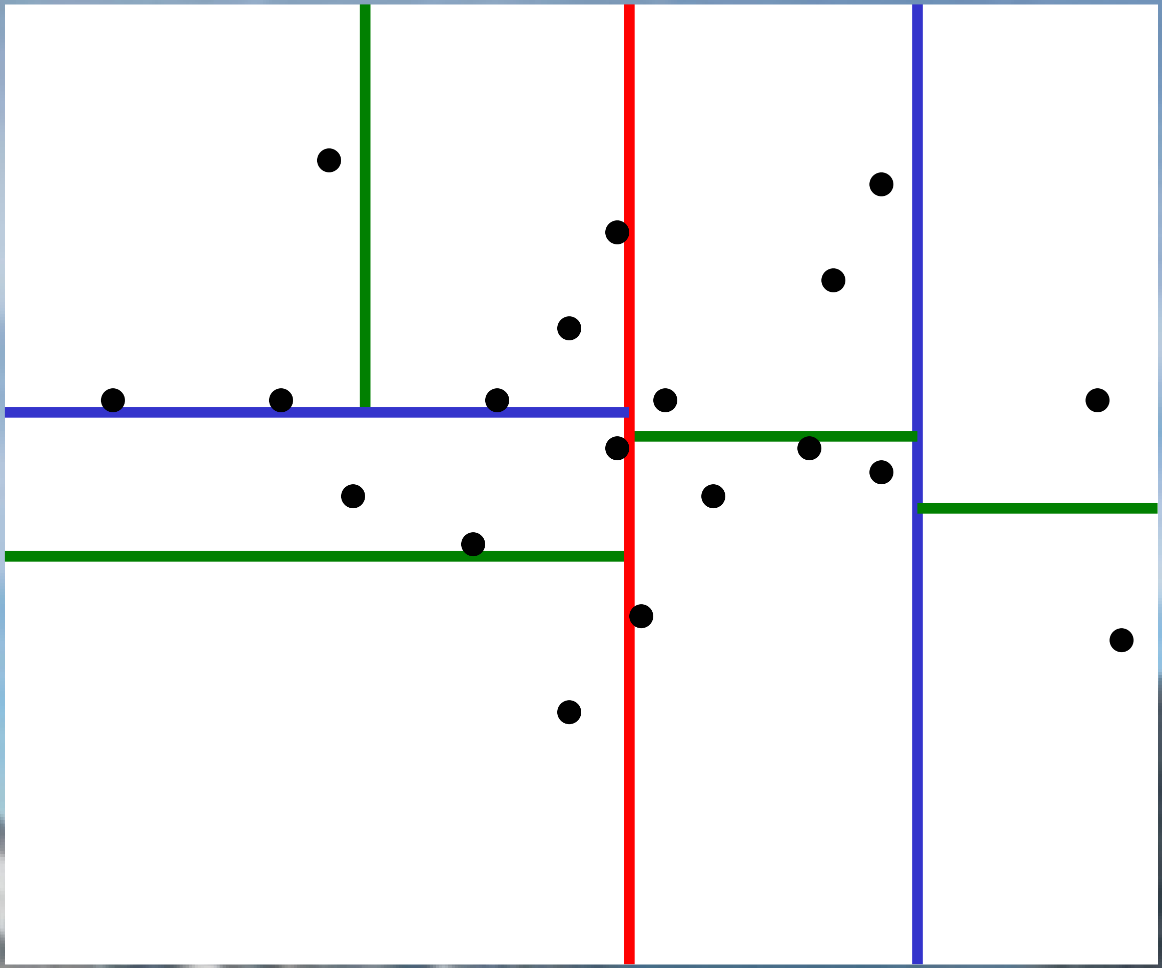
- Nearest neighbour matching

$$\forall j \; NN(j) = \arg\min_i ||\mathbf{x}_i - \mathbf{x}_j||, \; i \neq j$$

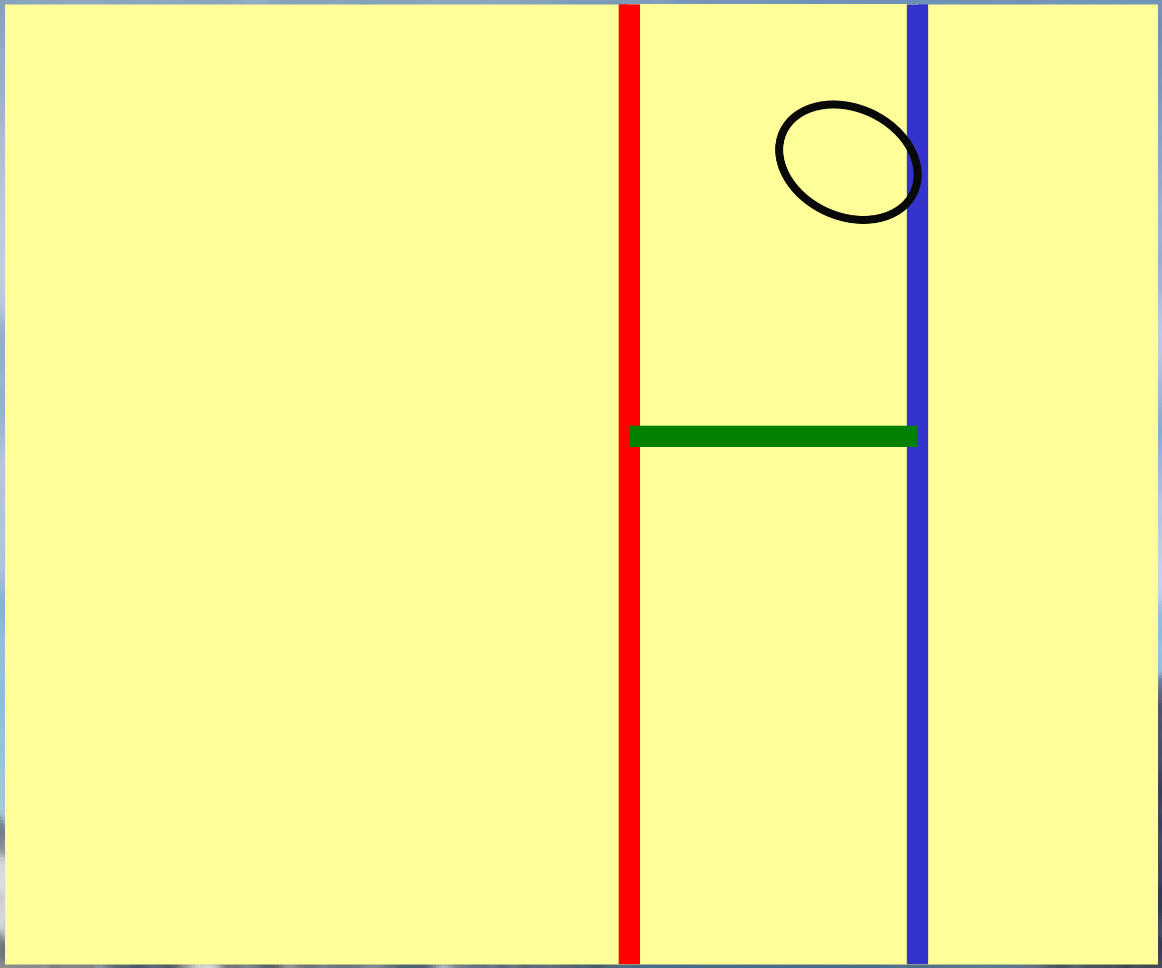
[Beis Lowe 1997, Nene Nayar 1997, Gray Moore 2000, Shakhnarovich 2003]

- Use k-d tree
  - k-d tree recursively bi-partitions data at mean in the dimension of maximum variance
  - Approximate nearest neighbours found in O(n log n)
- Find k-NN for each feature
  - k $\approx$ number of overlapping images (we use k = 4)

K-d tree



K-d tree



Overview

- Feature Matching
  - SIFT Features
  - Nearest Neighbour Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

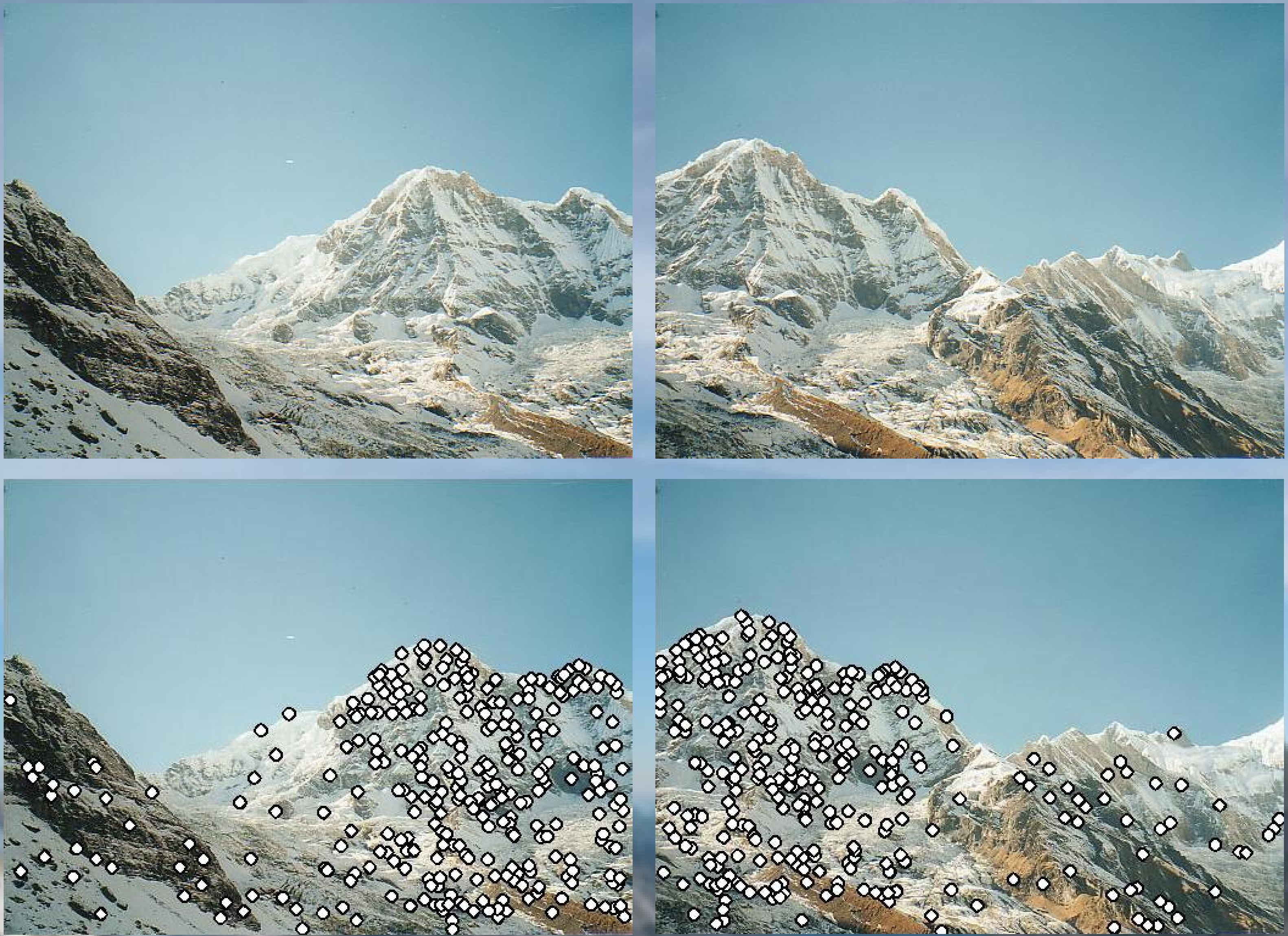
Overview

- Feature Matching
- Image Matching
  - RANSAC for Homography
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

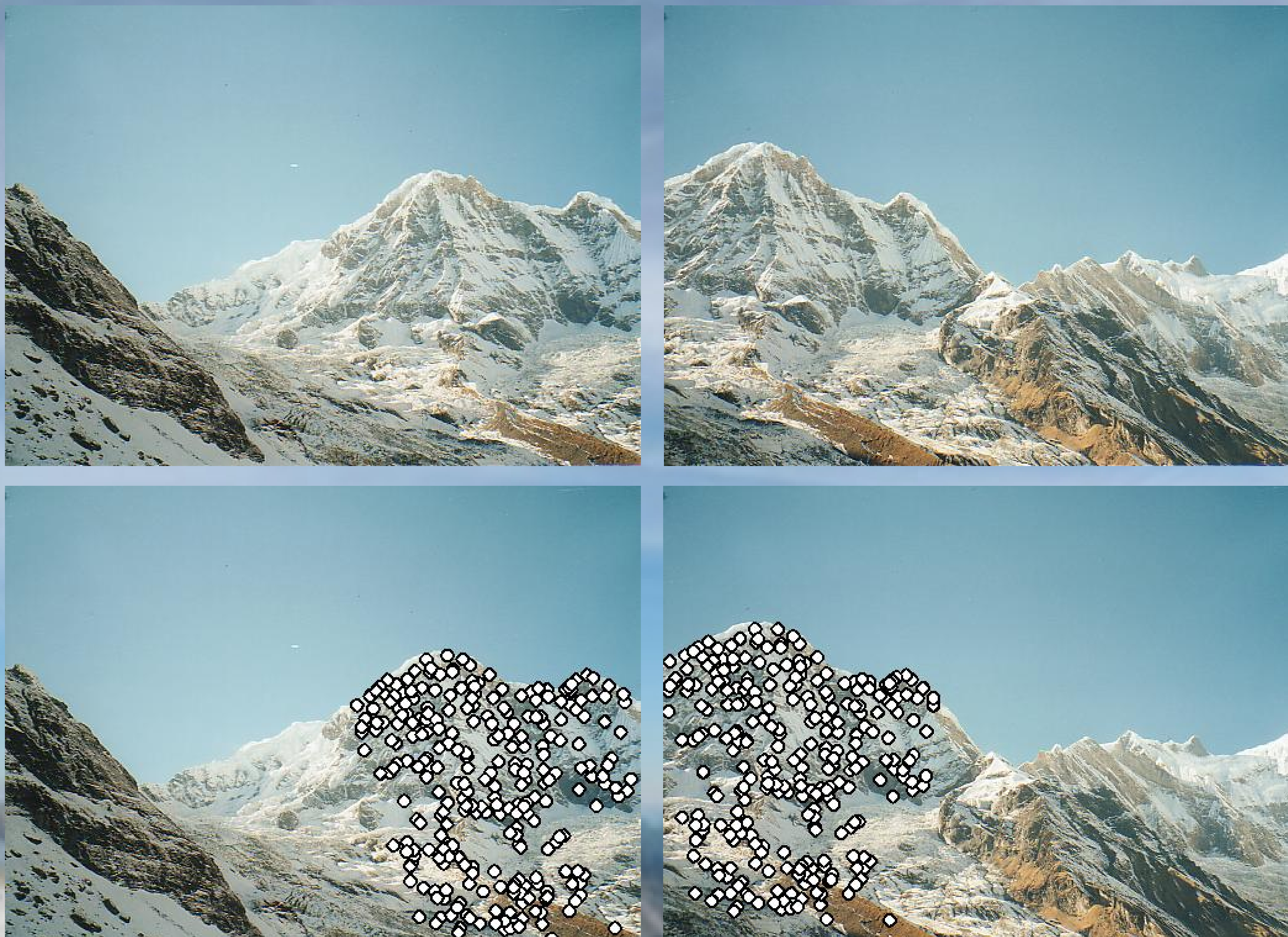
Overview

- Feature Matching
- Image Matching
  - RANSAC for Homography
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

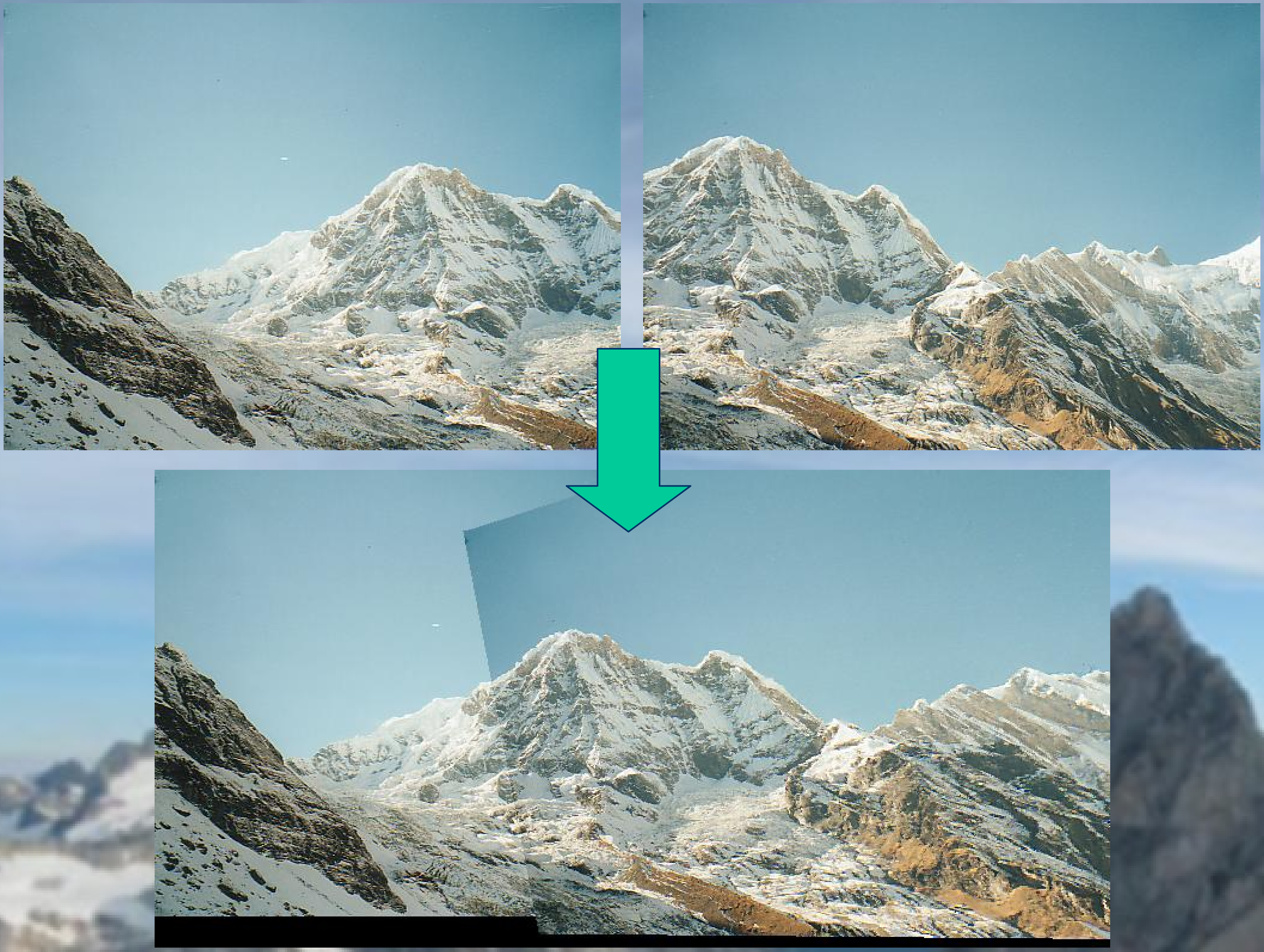
RANSAC for Homography



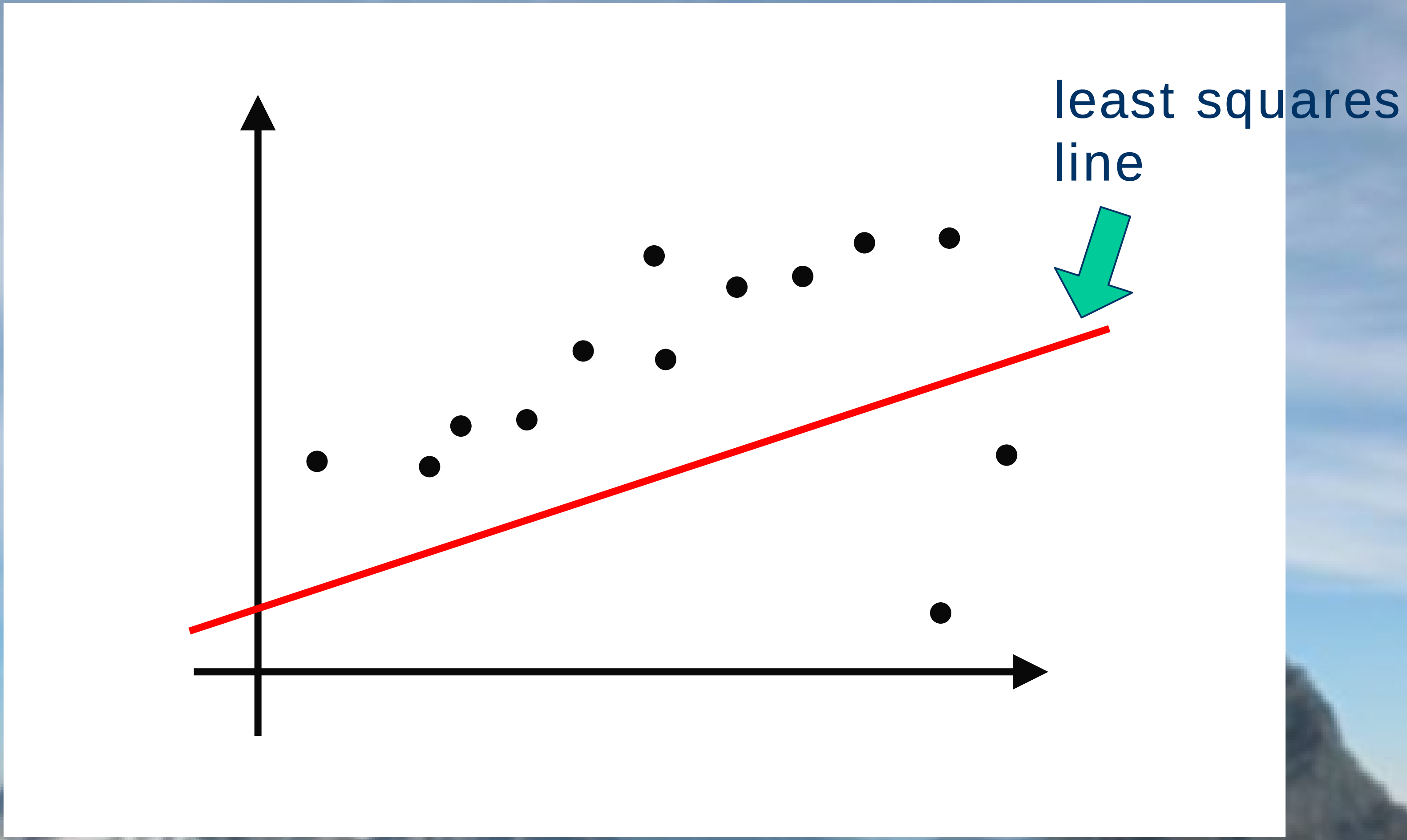
RANSAC for Homography



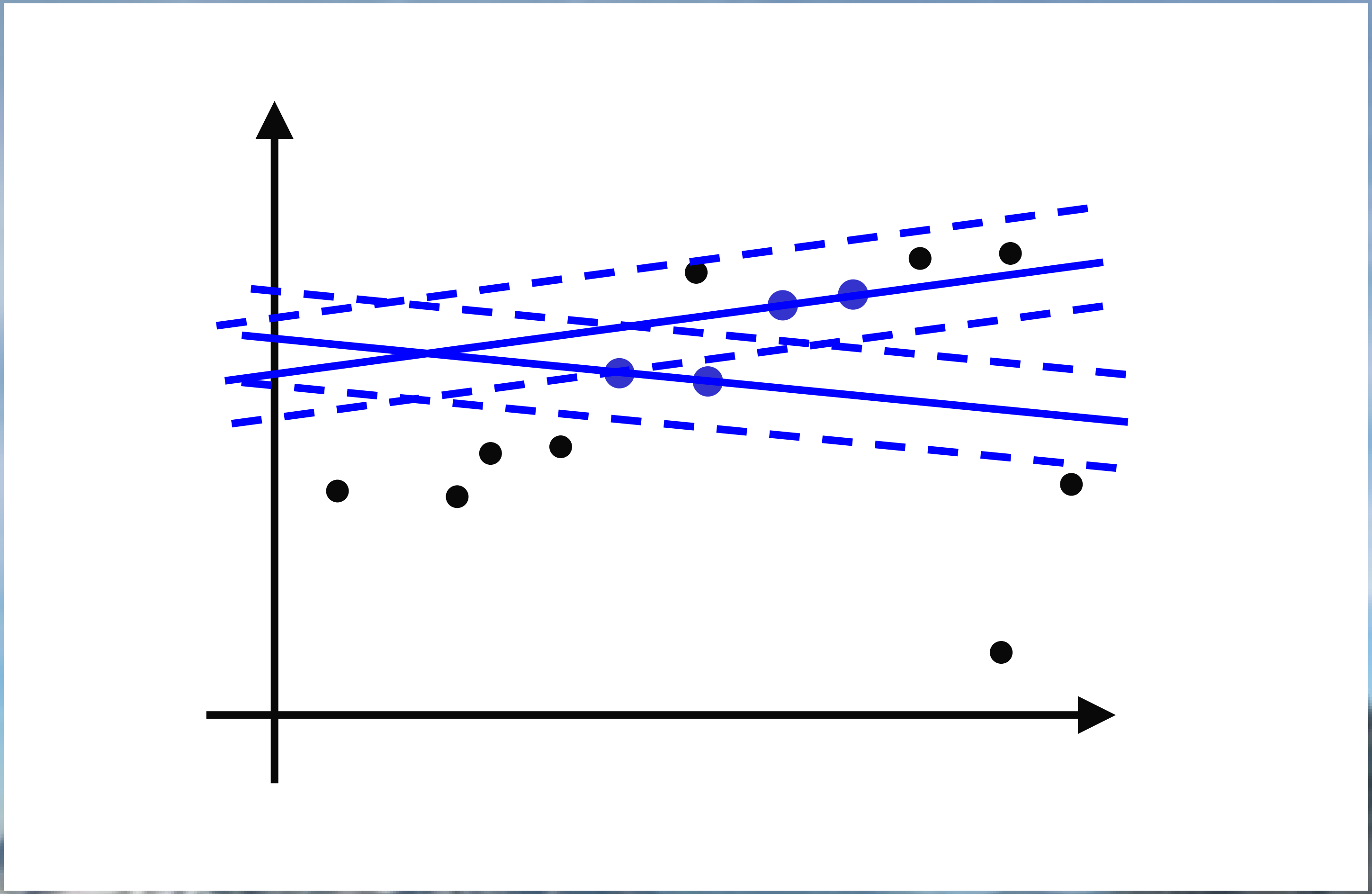
RANSAC for Homography



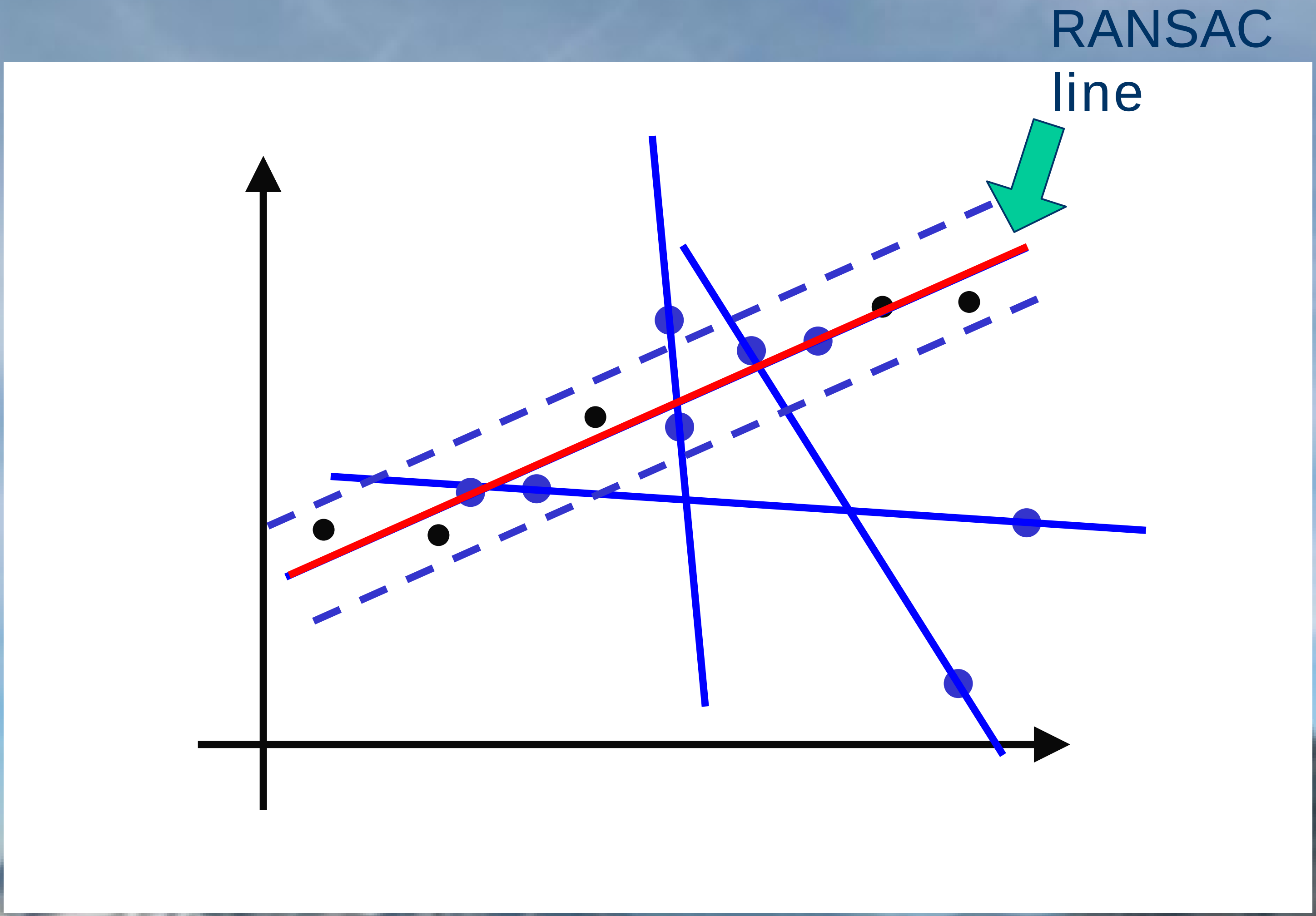
RANSAC: 1D Line Fitting



RANSAC: 1D Line Fitting



RANSAC: 1D Line Fitting



The RANSAC Algorithm

```
samplePoints = RandomSample(points);
```

2D Transforms

- Linear (affine)

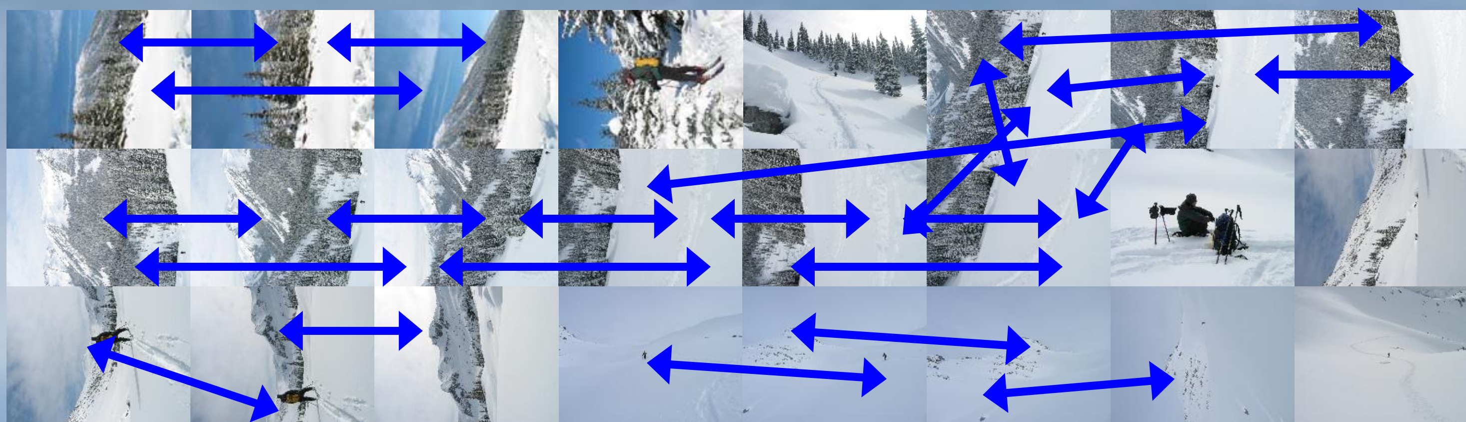
$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

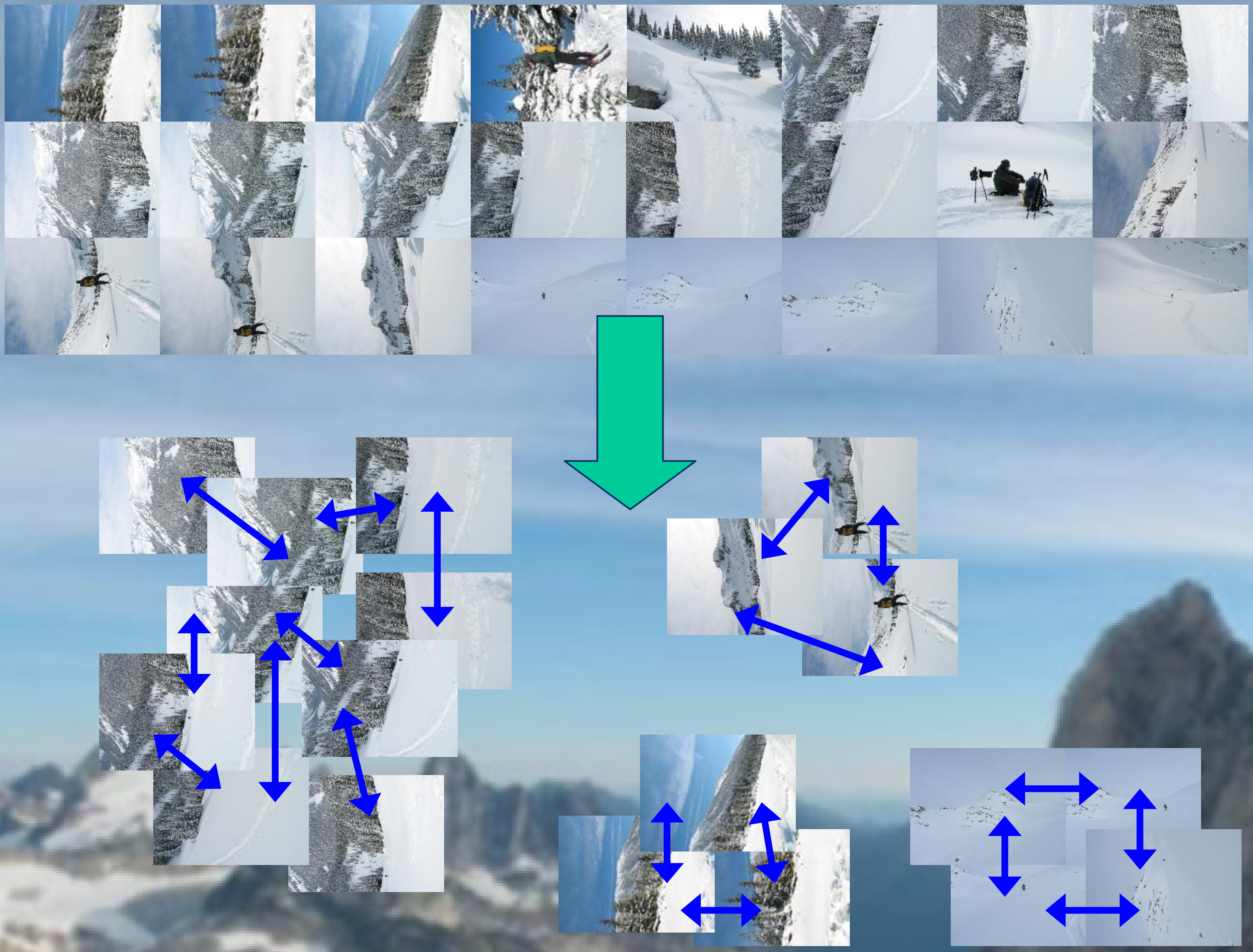
- Homography

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

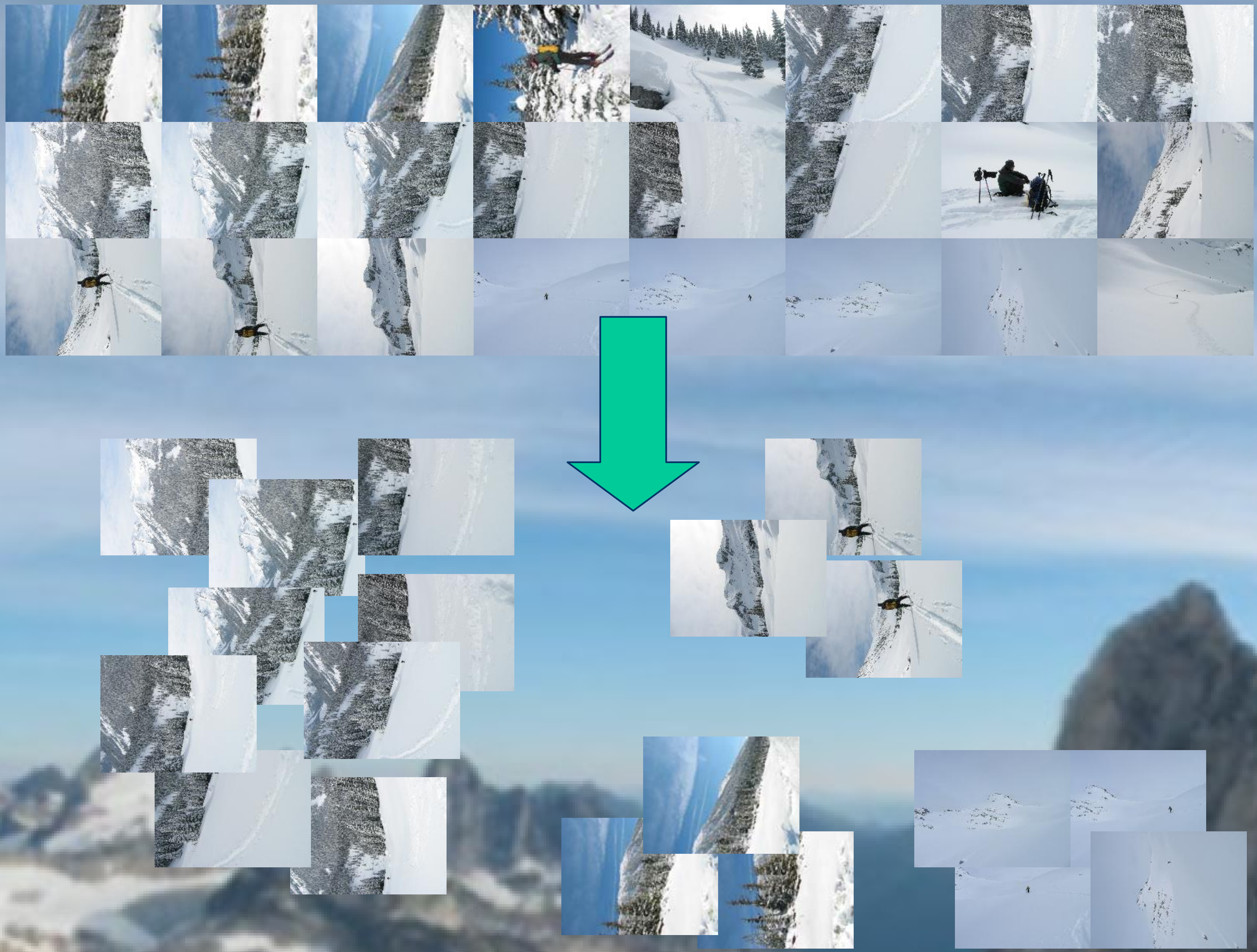
Finding the panoramas



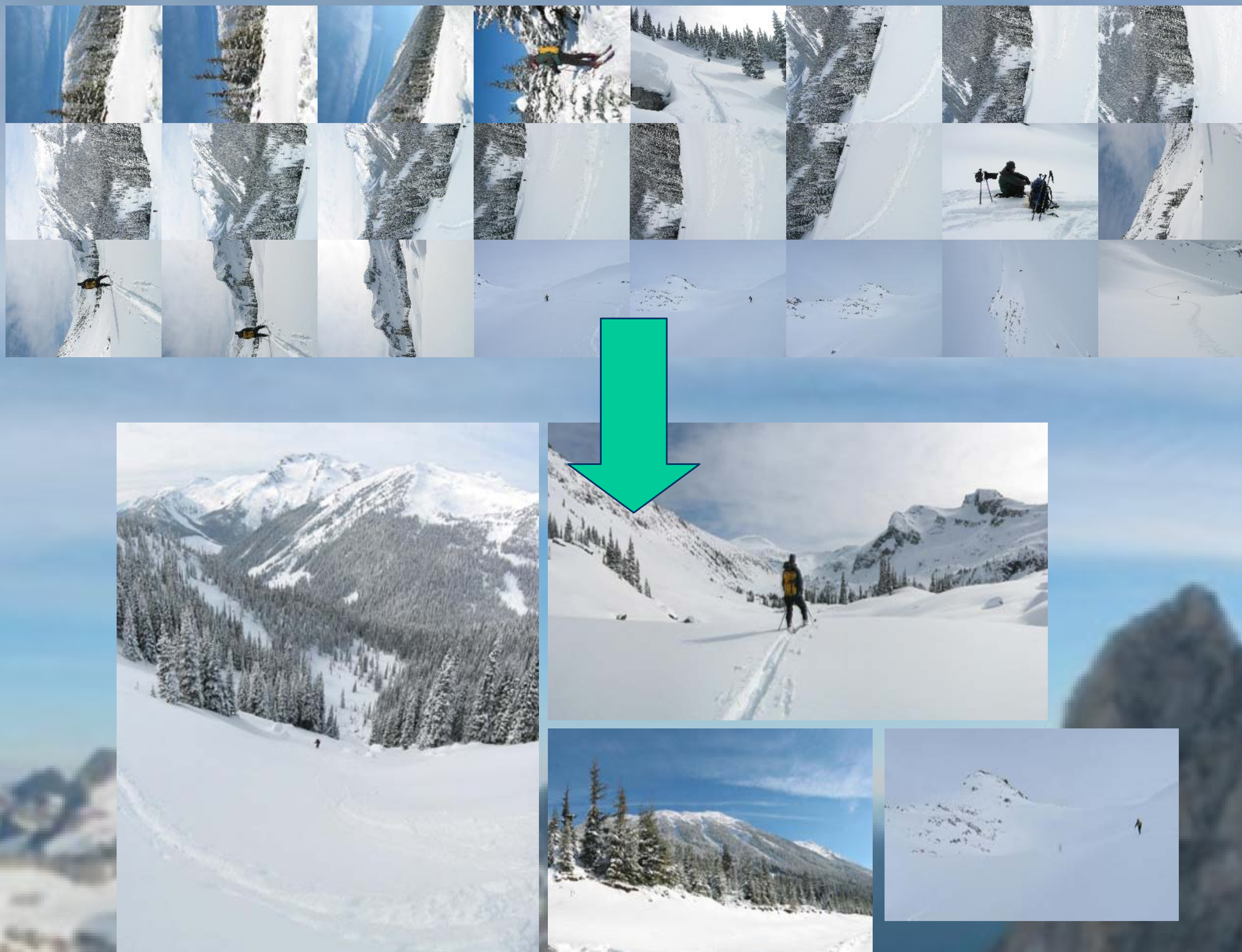
Finding the panoramas



Finding the panoramas



Finding the panoramas



Connected Components

```
% ConnectedComponent.m

% Find connected components in a square matrix
```

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Bundle Adjustment

- Adjust rotation, focal length of each image to minimise error in matched features

Bundle Adjustment

- Adjust rotation, focal length of each image to minimise error in matched features

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Multi-band Blending

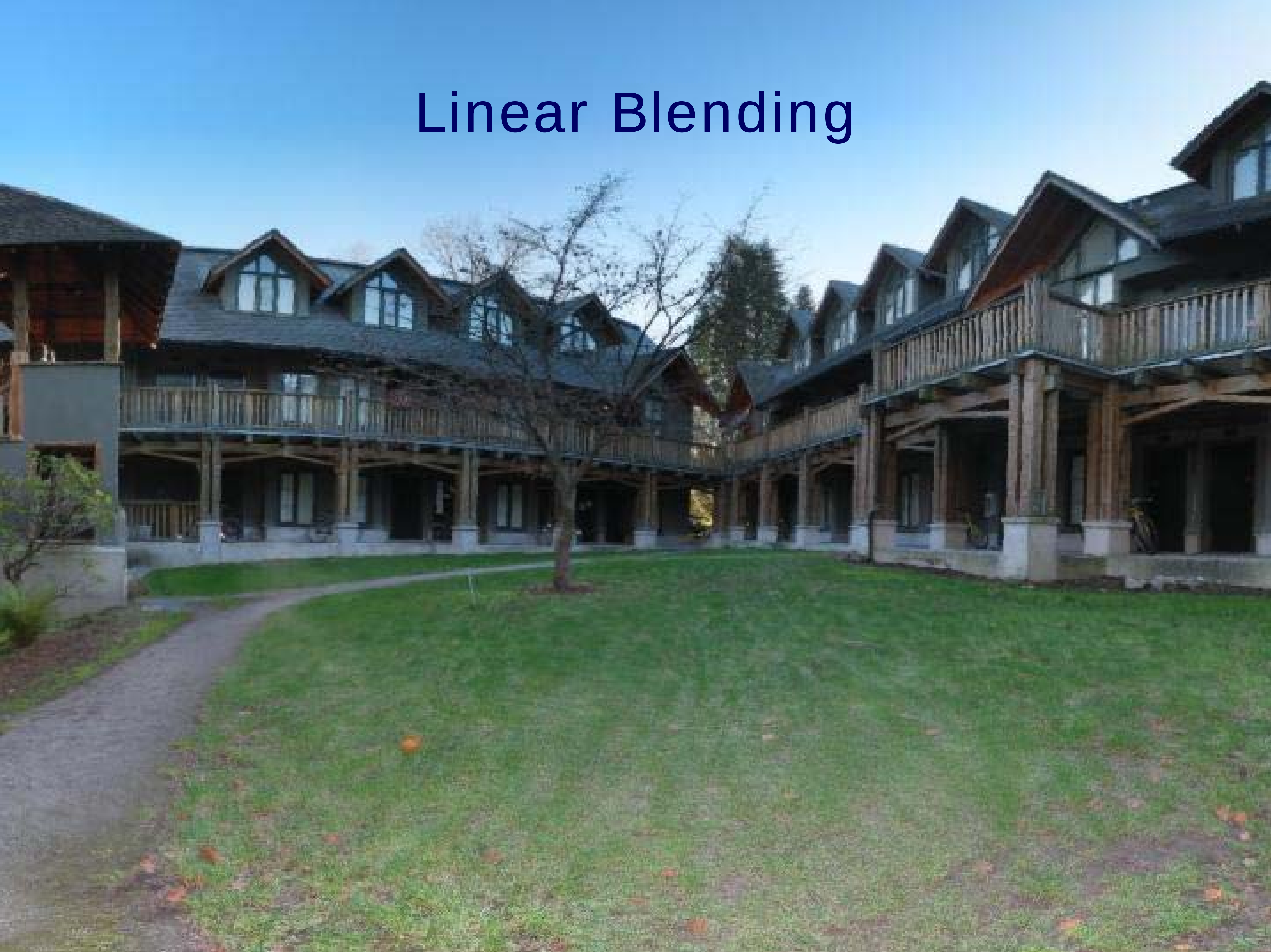
- Burt & Adelson 1983
  - Blend frequency bands over range $\propto \lambda$

2-band Blending

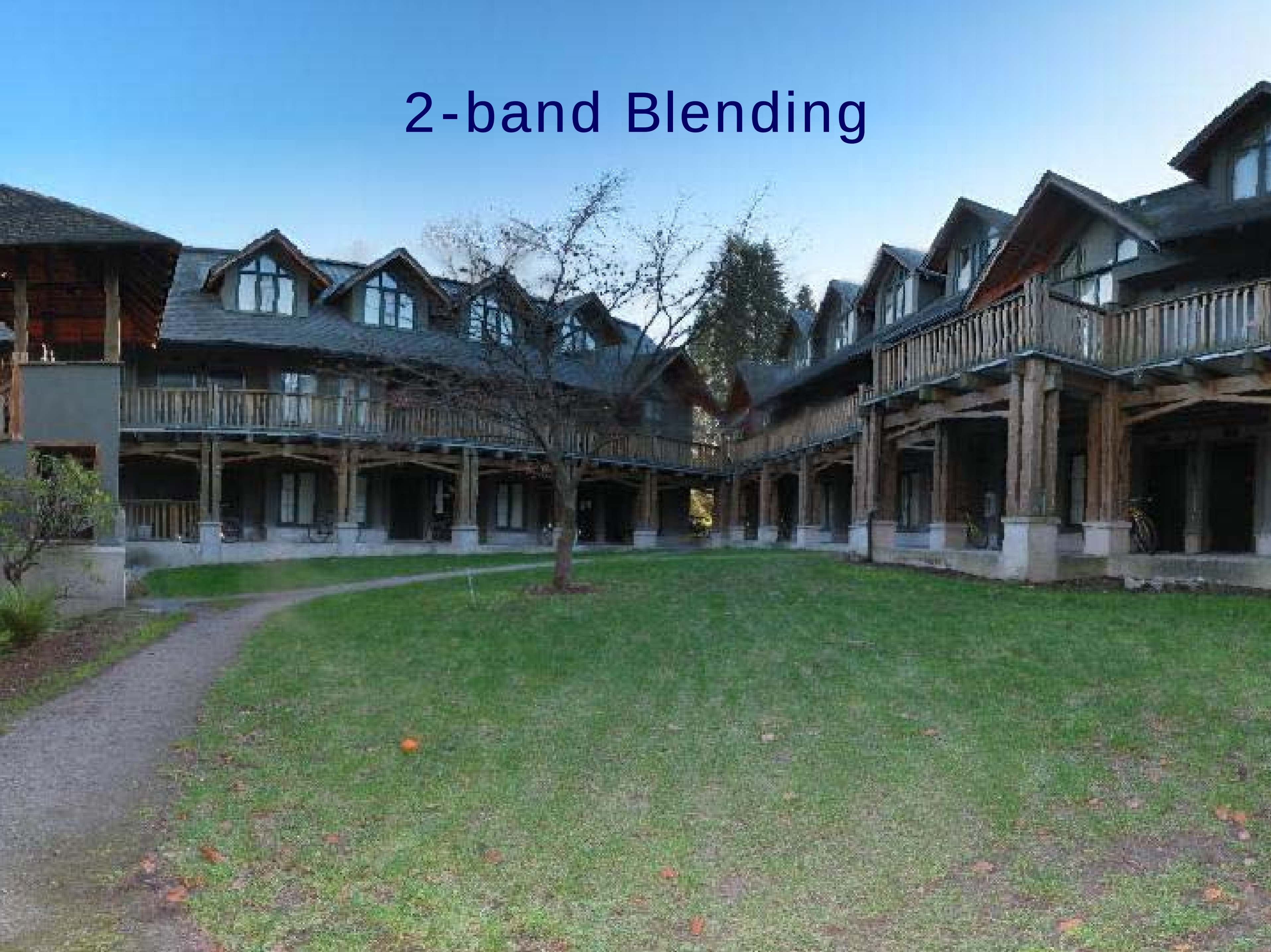
Low frequency ($\lambda > 2$ pixels)

High frequency ($\lambda < 2$ pixels)

Linear Blending



2-band Blending



Multi-band Blending

- No blending

Multi-band Blending

- Linear blending

  - Each pixel is a weighted sum

$$I^{linear} = \frac{\sum_i I^i W^i}{\sum_i W^i}$$

Multi-band Blending

- Multi-band blending

  - Each pixel is a weighted sum (for each band)

$$I_{k\sigma}^{multi} = \frac{\sum_i I_{k\sigma}^{i} \, W_{k\sigma}^{i}}{\sum_i W_{k\sigma}^{i}}$$

Multi-band Blending

- Linear blending

- Multi-band blending

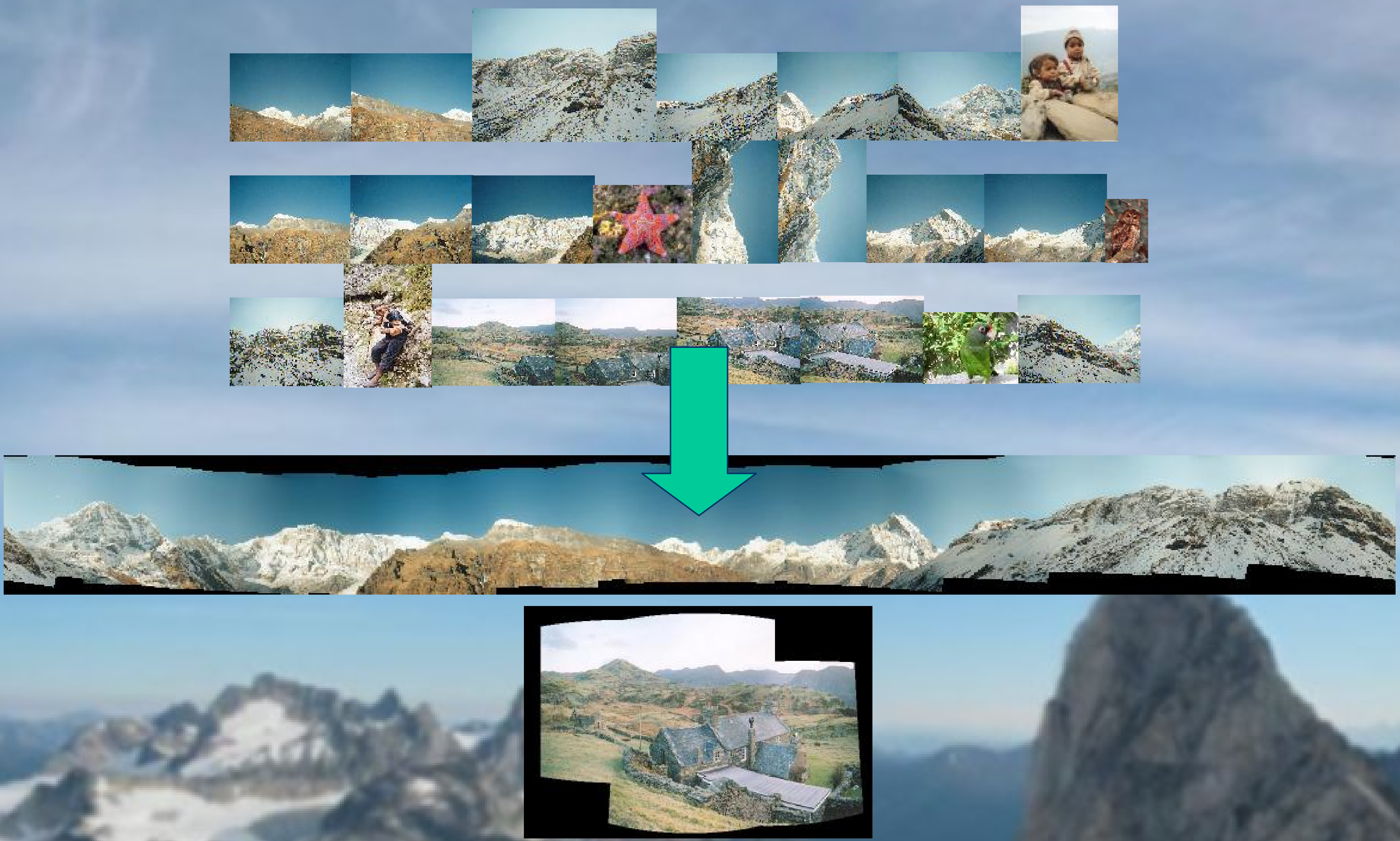
Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Results



Results

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions

Conclusions

- Fully automatic panoramas
  - A recognition problem…
- Invariant feature based method
  - SIFT features, RANSAC, Bundle Adjustment, Multi-band Blending
  - O(nlogn)
- Future Work
  - Advanced camera modelling
    - radial distortion, camera motion, scene motion, vignetting, exposure, high dynamic range, flash …
  - Full 3D case – recognising 3D objects/scenes in unordered datasets. “PhotoTourism”.

http://www.autostitch.net