Image Stitching
Based almost entirely on a presentation by Matthew Brown and David Lowe at ICCV 2003
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°
Human FOV

www.inition.co.uk/inition/guide_hmds_vrar.htm
Why “Recognising Panoramas”?

- 1D Rotations ($\theta$)
  - Ordering $\Rightarrow$ matching images
Why “Recognising Panoramas”?

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

- **2D Rotations** ($\theta, \phi$)
  - Ordering $\not\Rightarrow$ matching images
Why “Recognising Panoramas”?

• 2D Rotations ($\theta, \phi$)
  - Ordering $\neq$ matching images
Why “Recognising Panoramas”?

- 2D Rotations ($\theta, \phi$)
  - Ordering $\not\Rightarrow$ matching images
Why “Recognising Panoramas”? 
Overview

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

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

- Find k-NN for each feature
  - k ≈ number of overlapping images (we use k = 4)
- 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(nlogn)
Overview

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

- Feature Matching
- Image Matching
  - RANSAC for Homography
  - Finding the panoramas
- Bundle Adjustment
- Multi-band Blending
- Results
- Conclusions
Finding the panoramas
Finding the panoramas
Finding the panoramas
Finding the panoramas
Overview

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

- Sum of squared projection errors

\[ e = \sum_{i=1}^{n} \sum_{j \in I(i)} \sum_{k \in F(i,j)} f(r_{ij}^k)^2 \]

- \( n = \#\text{images} \)
- \( I(i) = \text{set of image matches to image } i \)
- \( F(i, j) = \text{set of feature matches between images } i, j \)
- \( r_{ij}^k = \text{residual of } k\text{th feature match between images } i, j \)

- Robust error function \( f(x) = \begin{cases} 
|x|, & \text{if } |x| < x_{max} \\
x_{max}, & \text{if } |x| \geq x_{max} 
\end{cases} \)
Homography for Rotation

- Parameterise each camera by rotation and focal length

\[ R_i = e^{[\theta_i]_x}, \quad [\theta_i]_x = \begin{bmatrix} 0 & -\theta_{i3} & \theta_{i2} \\ \theta_{i3} & 0 & -\theta_{i1} \\ -\theta_{i2} & \theta_{i1} & 0 \end{bmatrix} \]

\[ K_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

\[ \tilde{u}_i = H_{ij}\tilde{u}_j, \quad H_{ij} = K_iR_iR_j^TK_j^{-1} \]

- This gives pairwise homographies
Bundle Adjustment

- New images initialised with rotation,
Bundle Adjustment

- New images initialised with rotation,
Overview

- Feature Matching
- Image Matching
- Bundle Adjustment
- Warping
- Results
- Conclusions
Warping

- Take Notes
- Key
  - iterate in **target** image
  - calculate corresponding source pixel
  - resample
    - nearest neighbor (bad)
    - bilinear (better)
    - bicubic (best)
Overview

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

Burt & Adelson 1983
2-band Blending

Low frequency ($\lambda > 2$ pixels)

High frequency ($\lambda < 2$ pixels)
Linear Blending
2-band Blending
Overview

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

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

**Autopano: based on UBC’s code**

**live demo**