Large-scale 3D Modeling from Crowdsourced Data

Jan-Michael Frahm, UNC
Enrique Dunn, SIT
Jared Heinly, URCV
Johannes Schönberger, ETHZ
Image Registration and Connected-Component Discovery

Jared Heinly

URCV
urcvventures.com
Large-scale 3D Modeling from Crowdsourced Data
3D Modeling Pipeline

Data Association → Sparse Modeling → Dense Modeling
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
2D Features

• Represent an image by individual 2D components
2D Features

- Feature detection
- Decompose an image into salient points
2D Features

• Feature description
• Summarize points with distinctive representations

[4, 22, 5, 138, ...]
[19, 82, 71, 3, ...]
[105, 7, 48, 1, ...]
[54, 89, 6, 34, ...]
[96, 2, 122, 9, ...]
Feature Invariance

• Rotation
• Scale
• Affine
• Illumination
Rotation Invariance

Image Patch

Image Gradients
Rotation Invariance

Image Gradients

Gradient Histogram
Scale Invariance

• Image pyramid
Scale Invariance

- Image pyramid
- Find maximum response across scales
Affine Invariance
Illumination Invariance

• Descriptor normalization

\[ [4, 22, 5, 138, \ldots] \rightarrow [0.03, 0.15, 0.03, 0.92, \ldots] \]

\[ [19, 82, 71, 3, \ldots] \rightarrow [0.13, 0.55, 0.47, 0.02, \ldots] \]

\[ [105, 7, 48, 1, \ldots] \rightarrow [0.70, 0.05, 0.32, 0.01, \ldots] \]

\[ [54, 89, 6, 34, \ldots] \rightarrow [0.36, 0.59, 0.04, 0.23, \ldots] \]
Illumination Invariance

- Gradient direction

Image Gradients
2D Features

• Feature detection
• Decompose an image into salient points
2D Feature Detectors

• Blob
  • SIFT, SURF
2D Feature Detectors

- Blob

Scale

Gaussian

Difference of Gaussian
2D Feature Detectors

- Corner
  - Harris, FAST
2D Feature Detectors

• Corner
  • Harris, FAST
2D Feature Detectors

- Corner
  - Harris, FAST
2D Features

• Feature description

• Summarize points with distinctive representations

[4, 22, 5, 138, ...]

[19, 82, 71, 3, ...]

[105, 7, 48, 1, ...]

[54, 89, 6, 34, ...]

[96, 2, 122, 9, ...]
2D Feature Descriptors

- Gradients
  - SIFT, SURF

![Image Gradients]
2D Feature Descriptors

- Gradients
  - SIFT, SURF

\[ \begin{bmatrix} 0.012, & 0.492, & 0.187, & \ldots \end{bmatrix} \]

Feature Descriptor

Large-scale 3D Modeling from Crowdsourced Data
2D Feature Descriptors

- Binary comparisons
  - BRIEF, ORB, BRISK

Descriptor: 011010...
Outline – Data Association

• 2D features $\rightarrow$ Image registration
  • Features
  • Matching
  • Verification

• Image registration $\rightarrow$ Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Feature Matching
Feature Matching

• Distance metrics (real-valued)
  • Euclidean
  • Dot product

\[
\begin{bmatrix}
0.012, 0.492, 0.187, 0.618, 0.741, \\
0.913, 0.102, 0.003, 0.015, 0.120,
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.913, 0.102, 0.003, 0.015, 0.120, \\
\end{bmatrix}
\]
Feature Matching

- Distance metrics (binary)
  - Hamming

\[[0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, ...]\]

\[[0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, ...]\]
Feature Matching

- Which features to match
  - Exhaustive
Feature Matching

• Which features to match
  • Exhaustive
  • Approximate
Feature Matching
Feature Matching

• Match criteria
  • 1st nearest neighbor
  • Distance threshold

\[
\begin{bmatrix}
0.212, 0.492, 0.187, 0.618, \\
0.313, 0.102, 0.003, 0.515, \\
\vdots
\end{bmatrix}
\]

0.409
Feature Matching

• Match criteria
  • 1st nearest neighbor
  • Distance threshold
  • Ratio test

\[
\begin{align*}
&[0.212, 0.492, 0.187, 0.618, \ldots] \\
&[0.313, 0.102, 0.003, 0.515, \ldots] \\
&[0.602, 0.385, 0.661, 0.219, \ldots]
\end{align*}
\]

\[
\frac{0.409}{0.538} = 0.76
\]
Feature Matching

• Match criteria
  • Cross-check
  • One-to-one

[0.212, 0.492, 0.187, 0.618, ...]

[0.313, 0.102, 0.003, 0.515, ...]
Feature Matching
Feature Matching
Outline – Data Association

• **2D features → Image registration**
  - Features
  - Matching
  - Verification

• **Image registration → Connected components**
  - Exhaustive
  - Image retrieval
  - Clustering
  - Streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Epipolar Geometry
Epipolar Geometry

Epipolar Line
Feature Matching
Geometric Verification

• RANSAC (random sample consensus)
Geometric Verification

• RANSAC (random sample consensus)
Geometric Verification

• RANSAC (random sample consensus)
  • Solution with most inliers
Epipolar Geometry

- Camera intrinsics known
  - Essential matrix, $\mathbf{E}$ (5 points)
- Camera intrinsics unknown
  - Fundamental matrix, $\mathbf{F}$ (7 points)

$$
\mathbf{K} = \begin{bmatrix}
  f_x & c_x \\
  f_y & c_y \\
  1 & 1
\end{bmatrix}
$$

Epipolar Line
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

• 2D features $\rightarrow$ Image registration
  • Features
  • Matching
  • Verification

• Image registration $\rightarrow$ Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Connected Component Discovery

• How to discover related images?
• Exhaustive matching
• Nearest neighbors
  • Clustering and streaming are based off of this
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Exhaustive Matching

• Preemptive matching
  • Wu, 3DV 2013

• Determine candidate image pairs to verify
Exhaustive Matching

- Two-view geometry classification
  - Schönberger, CVPR 2015, GCPR 2015
- Determine candidate image pairs to verify
  - Image pair descriptor
Exhaustive Matching

- Two-view geometry classification
  - Schönberger, CVPR 2015, GCPR 2015
- Determine candidate image pairs to verify
  - Image pair descriptor

![Quantization Diagram]

- Feature location change
- Feature orientation change

![Prediction Diagram]

- Random forest classifier
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

- 2D features → Image registration
  - Features
  - Matching
  - Verification

- Image registration → Connected components
  - Exhaustive
  - Image retrieval
  - Clustering
  - Streaming
Image Retrieval

• Bag of words

[4, 22, 5, 138, ...]
[19, 82, 71, 3, ...]
[105, 7, 48, 1, ...]
[54, 89, 6, 34, ...]
[96, 2, 122, 9, ...]
Image Retrieval

- Bag of words
Image Retrieval

• Bag of words
Image Retrieval

- Bag of words
Image Retrieval

• VocMatch
  • Havlena and Schindler, ECCV 2014
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Clustering

- Frahm et al, ECCV 2010
- GIST
Clustering

• Frahm et al, ECCV 2010
• GIST
GIST Clustering

• Frahm et al, ECCV 2010
• GIST

![GIST Clustering](image)
Outline – Data Association

• 2D features $\rightarrow$ Image registration
  • Features
  • Matching
  • Verification

• Image registration $\rightarrow$ Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Streaming

- Limited memory budget
- Limited compute budget
- Handle very large datasets
Streaming

• Read images sequentially from disk
• Read each image only once
• Keep images in memory only as long as necessary
Challenges of Streaming

- Limited window in which to perform association
- No control over image order
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Image Registration

[Images of various landmarks and emojis]
Streaming Data Association

100M Images

Large-scale 3D Modeling from Crowdsourced Data
Streaming Data Association

100M Images
Streaming Data Association

100M Images

Large-scale 3D Modeling from Crowdsourced Data
Streaming Data Association

100M Images
Cluster Formation

100M Images

Iconic Image
Cluster Representation

Iconic Image

Bag of Visual Words

- 1632
- 82497
- 405
- 7189
- 94
Cluster Representation

Iconic Image

Bag of Visual Words

1632
82497
405
7189
94
Cluster Representation

Iconic Image

Bag of Visual Words

Cluster Image

Registered Visual Words

1632
82497
405
7189
94

1632
63917
383
7189
2219
Cluster Representation

Iconic Image

Bag of Visual Words

Cluster Image

Registered Visual Words

1632
82497
405
7189
94

1632
63917
383
7189
2219
Cluster Representation

Iconic Image

Bag of Visual Words
1632
82497
405
7189
94
63917
383
2219

Cluster Image

Registered Visual Words
1632
63917
383
7189
2219

Large-scale 3D Modeling from Crowdsourced Data
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Streaming Data Association

100M Images
A problem has been detected and Windows has been shut down to prevent damage to your computer.

DRIVER_IRQL_NOT_LESS_OR_EQUAL

If this is the first time you’ve seen this Stop error screen, restart your computer. If this screen appears again, follow these steps:

Check to make sure any new hardware or software is properly installed. If this is a new installation, ask your hardware or software manufacturer for any windows updates you might need.

If problems continue, disable or remove any newly installed hardware or software. Disable BIOS memory options such as caching or shadowing. If you need to use Safe Mode to remove or disable components, restart your computer, press F8 to select Advanced Startup Options, and then select Safe Mode.

**** ABCD.SYS - Address F73120AE base at C0000000, DateStamp 36B072A3

Kernel Debugger Using: COM2 (Port 0x2F8, Baud Rate 19200)
Beginning dump of physical memory
Physical memory dump complete. Contact your system administrator or technical support group.
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Cluster Discarding

Discard Rate

Large-scale 3D Modeling from Crowdsourced Data
Cluster Discarding
Cluster Discarding
Cluster Discarding
Memory Consumption

Number of Clusters in Memory

Number of Clusters (Thousands)

Number of Processed Images (Millions)

10K Discard Rate
Streaming Data Association

100M Images

Large-scale 3D Modeling from Crowdsourced Data
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Reconstruction Pipeline

- Data Association
- Sparse Modeling
- Dense Modeling

100M Images
Registered Images

1.5 Million Registered
105 Hours

Large-scale 3D Modeling from Crowdsourced Data

105
Scene Graph
Scene Graph

Strasbourg, France
Streaming Data Association

100M Images
Streaming Data Association

100M Images
Buckingham Palace, London, England
Golden Gate Bridge, San Francisco, California
Eiffel Tower, Paris, France
Taj Mahal, Agra, Uttar Pradesh, India
Other Connected Components
Other Connected Components
Other Connected Components
Other Connected Components
## Results

### Berlin, Germany (2.7M images)

<table>
<thead>
<tr>
<th></th>
<th>Frahm et al, 2010</th>
<th>Heinly et al, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered</td>
<td>4.6%</td>
<td>26%</td>
</tr>
<tr>
<td>Data Association</td>
<td>13.3 Hours</td>
<td>7.9 Hours</td>
</tr>
<tr>
<td>Time*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Identical Hardware Configuration
## Results

**Berlin, Germany (2.7M images)**

<table>
<thead>
<tr>
<th>Data Association Strategy</th>
<th>Match Attempts</th>
<th>Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frahm et al, 2010 GIST</td>
<td>1</td>
<td>4.6%</td>
</tr>
<tr>
<td><strong>Streaming</strong> GIST</td>
<td>2</td>
<td>8.9%</td>
</tr>
</tbody>
</table>
## Results

### Berlin, Germany (2.7M images)

<table>
<thead>
<tr>
<th>Data Association Strategy</th>
<th>Match Attempts</th>
<th>Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frahm et al, 2010 GIST</td>
<td>1</td>
<td>4.6%</td>
</tr>
<tr>
<td>Streaming GIST</td>
<td>2</td>
<td>8.9%</td>
</tr>
<tr>
<td>Streaming Bag-of-Words</td>
<td>1</td>
<td>25%</td>
</tr>
<tr>
<td>Streaming Bag-of-Words</td>
<td>2</td>
<td>26%</td>
</tr>
</tbody>
</table>
Challenges - Revisited

• Scalability
  • Streaming paradigm
  • Compact memory representation

• Completeness
  • Adaptive cluster representation
  • Favor diverse connections during streaming
Outline – Data Association

• 2D features → Image registration
  • Features
  • Matching
  • Verification

• Image registration → Connected components
  • Exhaustive
  • Image retrieval
  • Clustering
  • Streaming
Open-Source

github.com/jheinly

jaredheinly.com