Discovering Panoramas in Web Videos

Feng Liu¹, Yu-hen Hu² and Michael Gleicher¹

¹Department of Computer Sciences
²Department of Electrical and Comp. Engineering

University of Wisconsin-Madison
Discovering Panoramas in Web Videos

Feng Liu
fliu@cs.wisc.edu
How do you get a panoramic photo?
How do you get a panoramic photo?
A Two Step Process

Step 1: Find Appropriate Images

Step 2: Stitch them into a panorama
How do you get a panoramic photo?

A Two Step Process

Step 1: Find Appropriate Images

Step 2: Stitch them into a panorama

Easy! – A solved Computer Vision Problem!
• Lots of papers with good methods
• Common class project
• Good, available software (built into Windows, etc)
How do you get a panoramic photo?
A Two Step Process

Step 1: Find Appropriate Images

Step 2: Stitch them into a panorama

Easy! – A solved Computer Vision Problem!

- Lots of papers with good methods
- Common class project
- Good, available software (built into Windows, etc)

Requires appropriate images!
Videos as panorama sources

Videos often contain appropriate images
Segments of video contain panorama sources

Not all video is appropriate
Discover Panoramas in Video

Find good panorama segments inside longer videos (and make panoramas from them)
An Application

Discover Panoramas from Web Queries

Good panorama segments inside longer videos inside web libraries
Road Map

• Introduction
• Background
• Metrics for Panoramas from Video
• Finding Video Segments for Panoramas
• Panorama Assembly Tricks
• Results
Panorama Building: History

Along the River During Ching Ming Festival
by Z.D Zhang (1085-1145)

San Francisco from Rincon Hill, 1851, by Martin Behrmanx
Panorama Building: A Concise History

The state of the art and practice is good at assembling images into panoramas

- Mid 90s – Commercial Players (e.g. QuicktimeVR)
- Late 90s – Robust stitchers (in research)
- Early 00s – Consumer stitching common
- Mid 00s – Automation
Stitching Recipe

• Align pairs of images

• Align all to a common frame

• Adjust (Global) & Blend
Making a Digital Panorama
Stitching Images Together
Background

When do two images “stitch”? 

Images taken from the same viewpoint are related
Images can be transformed to match
Images related by *Homographies*

- 8 parameter, 2D Image Transformation

\[
x', y' = \frac{ax + by + c}{gx + hy + 1}, \frac{dx + ey + f}{gx + hy + 1}
\]

\[
\begin{bmatrix}
w x' \\
w y' \\
w
\end{bmatrix} = \begin{bmatrix}
a & b & c \\
d & e & f \\
g & h & i
\end{bmatrix} \begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]
Finding Homographies

- Find Corresponding Features*
- Compute Best-Fit Homography (using robust statistics e.g. RANSAC)

Two images stitch if and only if the best fit homography is a good fit
- If the best fit homography is a bad fit, the resulting panorama will be bad.
Automatic Feature Point Matching

- Match local neighborhoods around points
- Use descriptors to efficiently compare
  - SIFT [Lowe98] most common choice
Automatic Recognition of Panoramas in Sets of Still Images

### Similar problems: Hard for different reasons

<table>
<thead>
<tr>
<th>Recognizing Panoramas in Sets of Stills</th>
<th>Discovering Panoramas in Web Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Not all images work</td>
<td>• Not all videos work</td>
</tr>
<tr>
<td>• Matching is Hard</td>
<td>• Matching is not so hard</td>
</tr>
<tr>
<td>• Images unordered</td>
<td>• Images are ordered</td>
</tr>
<tr>
<td>• Large differences in images</td>
<td>• Continuity limits differences</td>
</tr>
<tr>
<td>• Different Orientations</td>
<td>• Orientations consistent</td>
</tr>
<tr>
<td>• High Quality Images</td>
<td>• Variable Image Quality</td>
</tr>
<tr>
<td>• Relatively small image sets</td>
<td>• Potentially large image sets</td>
</tr>
<tr>
<td>• Assume sufficient coverage</td>
<td>• Small motions uninteresting</td>
</tr>
<tr>
<td></td>
<td>• Dynamic Objects</td>
</tr>
</tbody>
</table>
Road Map

- Introduction
- Background
- Metrics for Panoramas from Video
- Finding Video Segments for Panoramas
- Panorama Assembly Tricks
- Results
When does a segment of video lead to a good panorama?

- Good homographies between frames
- Individual images have good quality
- Result has a wide field of view

Competing interests:
- More frames give wider field of view
- More frames give more accumulated error
Homography Quality

- Points should match (robust best fit)
- Measure residual distances
- Penalize large residuals
Source Image Quality

Blurriness=0.049

Blurriness=0.299

Method of [Tong et al 04]

Blockiness = .204

Blockiness=.497

Method of [Wang et al 02]
Field of View (Coverage)

- Pick the base frame that minimizes area
- And therefore has minimum distortion
Good Segments

- Images fit together well
  - Pair-wise homography residuals small
- Images are of high quality
  - Each per-frame quality penalty is small
- Covers a wide field of view
  - Minimum area covered is large
An Optimization Problem

Given a Video V
Find (non-overlapping) segments $S_i$ that
- Have maximal field of view / coverage
- Have minimal penalties
  - Homography error
  - Image quality penalty

Reject segments that have
- Too little coverage
- Too much penalty
In Practice

Greedy, Approximate Algorithm

Detect shot boundaries

Divide segments that have too much penalty
Repeat until done

Discard segments with too little coverage
Example Results

(a) West lake 1. Its source images are frame 0 to 86 from a 26-second, 10fps web video.

(b) West lake 2. Its source images are frame 89 to 251 from the same video as (a).

(c) Lake Louise. Its source images are frame 0 to 98 from a 21-second, 10fps web video.
More Examples

(d) Arches National Park. Its source images are frame 55 to 173 from a 15-second, 15fps web video.

(e) Arches National Park. Its source images are frame 142 to 265 from a 31-second, 15fps web video.

(f) Arches National Park. Its source images are frame 1168 to 1296 from a 83-second, 15fps web video.
What about moving objects?

Detect [Liu&Gleicher06]

Discard

Selectively add back in
Activity Synopsis Examples

(a) Horse riding. Its source images are frame 207 to 495 from a 58-second, 15fps web video.

(b) Biking. The source images of the two panoramas are frame 2729 to 2856, and frame 4298 to 4625 from a 194-second 30fps web video.
Variable Image Quality

- Images in video have varying image quality
  - Compression artifacts, motion blur, ...

Weight blends by image quality.
Bad images contribute less
Usage Scenario
Panorama from Web Query

- Query YouTube for a search term
- Fetch top 10 hits*
- Try panorama discovery in each

* Presently done manually
Evaluation

- Tried 6 queries (60 total videos)
- Created panoramas from most (87%)

- Compare discovery with human expert
  - Expert only looks for camera motions
  - Algorithm looks for panorama sources

- Never found panoramas that expert did not
- Found 87% of those identified by expert
Limitations

Predicts success of *our* stitcher

- Only considers linear ordering
  - Can’t connect 2D layouts
  - Can’t skip over bad frames

- Only planar panoramas
More examples
A final example
90 seconds of seagulls?

YouTube Query: Vancouver Beach
You can find good (ok, decent) panoramas in surprising web videos

From YouTube query: “Vancouver Beach”
Find video segments that yield panoramas:
• Fit the panoramic (homography) model
• Have good image quality
• Create wide field-of-view panoramas

This work was supported in part by NSF grant IIS-0416284.
Videos from YouTube and Image Examples from Flickr & Panorama Websites
Background: Panorama Assembly

Register all images to a common “base frame”

Blend to hide seams

Images from Brown & Lowe ‘03