Scene Completion
Using Millions of Photographs

James Hays and Alexei A. Efros
Carnegie Mellon University

Thanks to James & Alyosha for slides!

Fill in unknown region from source image parts

Hays and Efros, SIGGRAPH 2007
Efros and Leung result – no notion of semantics, also assumes necessary data is present elsewhere in the image.
Scene Matching for Image Completion
Scene Completion Result

Challenges:

Computational costs of searching lots of images

Should fill in missing regions with semantically valid fragments
The Algorithm

- Input image
- Scene Descriptor
- Image Collection

20 completions

Context matching + blending

200 matches

Hays and Efros, SIGGRAPH 2007
Data

We downloaded **2.3 Million** unique images from Flickr groups and keyword searches.

Groups: lonelyplanet, urban-fragments, ruraldecay ...
Keywords: outdoors, vacation, river...
Discard duplicates and small images
Scene Matching

Hays and Efros, SIGGRAPH 2007
Scene Descriptor

Compute oriented edge response at multiple scales (5 spatial scales, 6 orientations)

Hays and Efros, SIGGRAPH 2007
Scene Descriptor

Gist scene descriptor (Oliva and Torralba 2001)
“semantic” descriptor of image composition
aggregated edge responses over 4x4 windows
scenes tend to be semantically similar under this descriptor if very close

Hays and Efros, SIGGRAPH 2007
Scene Descriptor

Gist scene descriptor - with missing regions masked (weighted based on percentage of valid pixels)

Hays and Efros, SIGGRAPH 2007
Scene Descriptor

Color descriptor – color of the query image downsampling to 4x4

Distances calculated by SSD between query image descriptors & imgs in database

Total Dist = color dist + 2 * gist dist

Gist scene descriptor (Oliva and Torralba 2001)

Hays and Efros, SIGGRAPH 2007
Need to more precisely align matching scenes to local img context around missing region
local context = all pixels within 80 pixel radius of hole’s boundary
Compute pixel-wise error of 200 best scene matches across all valid translations and 3 scales
Compute texture similarity of proposed fill-in to removed region

Hays and Efros, SIGGRAPH 2007
Final result – blended between the two images along the cut to merge seamlessly

Hays and Efros, SIGGRAPH 2007
We assign each of the 200 results a score which is the sum of:

1. The scene matching distance
2. The context matching distance (color + texture)
3. The graph cut cost

Hays and Efros, SIGGRAPH 2007
Top 20 Results

Hays and Efros, SIGGRAPH 2007
Pro – allows insertion of novel objects
The Algorithm

Input image
The Algorithm

Input image  Scene Descriptor

Compute a global description of the whole image

Hays and Efros, SIGGRAPH 2007
The Algorithm

Input image  ➔ Scene Descriptor  ➔ Image Collection

Compare to LOTS of images

Hays and Efros, SIGGRAPH 2007
The Algorithm

Input image → Scene Descriptor → Image Collection

Get top matches

200 matches

Hays and Efros, SIGGRAPH 2007
The Algorithm

Input image ➔ Scene Descriptor ➔ Image Collection

Compare more locally and merge pieces of matching images

Context matching + blending

200 matches
Hays and Efros, SIGGRAPH 2007
The Algorithm

Input image

Scene Descriptor

Image Collection

20 completions (final results)

Context matching + blending

200 matches

Hays and Efros, SIGGRAPH 2007
200 scene matches

Hays and Efros, SIGGRAPH 2007
... 200 scene matches

Hays and Efros, SIGGRAPH 2007
... 200 scene matches

Hays and Efros, SIGGRAPH 2007
... 200 scene matches
... 200 scene matches
... 200 scene matches

Hays and Efros, SIGGRAPH 2007
... 200 scene matches

Hays and Efros, SIGGRAPH 2007
Failures
Failures
Cause of failure – atypical scene caused lack of good matches
Failures

Hays and Efros, SIGGRAPH 2007
Failures

Hays and Efros, SIGGRAPH 2007
Failures

Hays and Efros, SIGGRAPH 2007
Failures

Cause of failure – fine scale texture mismatch

Hays and Efros, SIGGRAPH 2007
Failures
Failures

Hays and Efros, SIGGRAPH 2007
Failures

Hays and Efros, SIGGRAPH 2007
Failures

Cause of failure – no notion of “objects”
Evaluation
Real Image. This image has not been manipulated

or

Fake Image. This image has been manipulated

Hays and Efros, SIGGRAPH 2007
User Study Results - 20 Participants

![Graph showing the percentage of images marked fake against maximum response time (seconds). Three curves are shown: Criminisi et al., Our algorithm, and Real Photographs. The graph is sourced from Hays and Efros, SIGGRAPH 2007.]
Why does it work?
10 nearest neighbors from a collection of 20,000 images

Hays and Efros, SIGGRAPH 2007
10 nearest neighbors from a collection of 2 million images

Hays and Efros, SIGGRAPH 2007
The Big Picture

Brute-force Image Understanding – insert semantically matching pieces by looking through millions of images.

Hays and Efros, SIGGRAPH 2007
The Big Picture

Brute-force Image Understanding – classify an image by looking at millions of other images

Torralba, Fergus, and Freeman, 80 Million Tiny Images.

Hays and Efros, SIGGRAPH 2007