Automatic Photo Popup: 3D Reconstruction from Single Image

Presented by

Santosh Kumar Divvala
Carnegie Mellon University
The World Behind the Image
The Problem

- Recovering 3D geometry from **single** 2D projection
- Infinite number of possible solutions!

from [Sinha and Adelson 1993]
Our World is Structured

Abstract World

Our World

Image Credit (left): F. Cunin and M.J. Sailor, UCSD
Infer Most Likely Scene

Unlikely

Likely
Our Goals

• Simple, piecewise planar models
• Outdoor scenes
• Doesn’t need to work all the time (~35%)
Approach: Learning

- Learn structure of the world and appearance-based models of geometry
More samples from our dataset
Geometric Cues

- Color
- Texture
- Location
- Perspective
Get Good Spatial Support

50x50 Patch

50x50 Patch
Image Segmentation

- Simple Idea #1: segment the image – Won’t work!
- Chicken & Egg problem
- Simple Idea #2: multiple segmentations
  - Decide later which segments are good
Labeling Segments

For each segment:

- Get $P(\text{good segment} \mid \text{data}) \ P(\text{label} \mid \text{good segment, data})$
Image Labeling

Labeled Segmentations

Labeled Pixels

\[ P(label|data) \propto \sum_{segments} P(good\ segment|data)P(label|good\ segment, data) \]
Cutting and Folding

- Fit ground-vertical boundary
  - Iterative Hough transform
Cutting and Folding

- Construct 3D model
- Texture map
Summary

Input  Geometric Labels  Cut’n’Fold  3D Model

Image

Ground

Vertical

Sky

Learned Models
Results

Input Image

Cut and Fold

Automatic Photo Pop-up
Results

Input Image

Cut and Fold

Automatic Photo Pop-up
Results

Input Image

Automatic Photo Pop-up
Results

Input Images

Automatic Photo Pop-up
Results

Input Image

Automatic Photo Pop-up
Comparison with Manual Method

Input Image

Automatic Photo Pop-up (30 sec)!

[Liebowitz et al. 1999]
Surface Estimates: Indoor

Avg. Accuracy
Main Class: 93%
Subclass: 76%
Failures

Labeling Errors
Failures

Foreground Objects
Applications: Robot Navigation

[Nabbe Hoiem Hebert Efros 2006]
Applications: Object Detection

- Object detector:
- Learn distribution of object heights *in the 3D world*
Object Detection Result

Car: TP / FP  Ped: TP / FP

Initial: 2 TP / 3 FP
Final: 7 TP / 4 FP
The Music Video
Its on the Web!

Get Inside Your Picture
What if you could transform 2D into 3D?

At FotoWoosh we believe that the greatest and most powerful ideas have been expressed on a page. The more we can transform what can live on a page the more powerful and impactful everyone’s ideas can be expressed. FotoWoosh is one of our efforts to transform what is possible on a page. We want to revolutionize what you think a website is and what can live on a website. FotoWoosh will forever change the way you think of your pictures. We live in a multidimensional world; it’s time to break out of the one or two dimensions that media and technology have locked us into for so long. Let’s dive in and look around. Let’s create worlds where there once were barriers. Learn more →

view samples
Thank you

Questions?