Scene Shape from Texture of Objects

Nadia Payet and Sinisa Todorovic

Oregon State University OSU

CVPR 2011 Colorado Springs

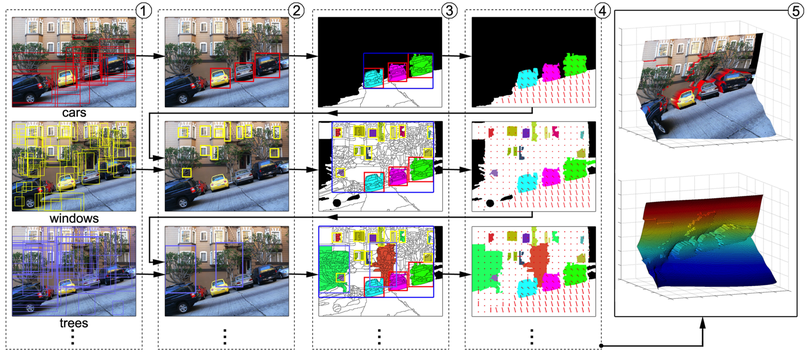
PROBLEM STATEMENT

Given noisy image detections

Identify the right detections and extract their texture

Reconstruct the 3D shape

OUR APPROACH



1. Extract noisy object detections with the detector of [1]
2. Label the right detections sequentially $\Rightarrow$ texture elements
3. Estimate normals for each detection
4. Diffuse the normals for shape from texture
5. Reconstruct the remaining scene parts

DETECTING TEXTURE

SEARN [2] sequentially labels detections as texture elements

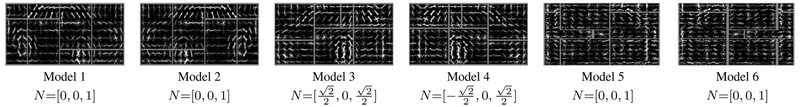
Apply classifier f to each detected bounding box

$$f^{(\tau+1)} = \beta h^{(\tau+1)} + (1-\beta) f^{(\tau)}$$

$h^{(\tau+1)}$: weak classifier

Accumulate contextual information to help remove false positives

SHAPE FROM TEXTURE



Model 1 $N=[0,0,1]$ | Model 2 $N=[0,0,1]$ | Model 3 $N=[\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}]$ | Model 4 $N=[-\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}]$ | Model 5 $N=[0,0,1]$ | Model 6 $N=[0,0,1]$

We compute surface normals of objects from their canonical poses

Surface normal of object $i \rightarrow N_i = H_i N \leftarrow$ Model normal

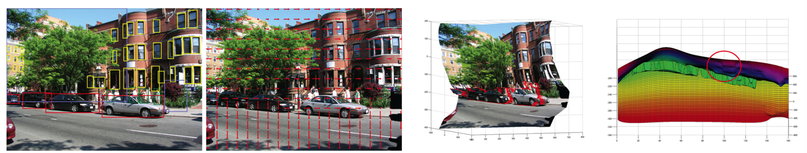
H_i: Homography from model parts to detected parts

Diffusion of object normals $\Rightarrow$ Shape of texture surfaces

Domain knowledge $\Rightarrow$ Shape of non-texture surfaces

[1] Felzenzwalb et al., Object detection with discriminatively trained part-based models, PAMI 2010.
[2] Daume et al., Search-based structured prediction, Machine Learning Journal 2009.
[3] Gupta et al., Blocks world revisited, ECCV 2010.

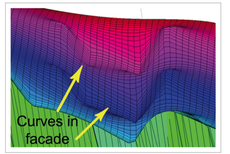
RESULTS



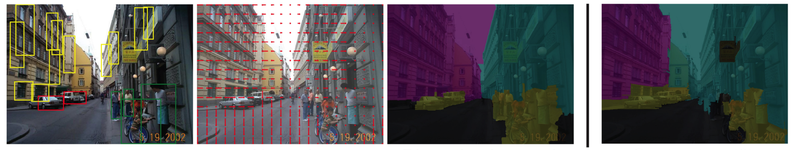
Scene reconstruction for the LabelMe dataset

Method	Car	Window	Tree	Ped.	All
[1] low threshold	0.574	0.418	0.521	0.592	0.526
[1] learned threshold	0.812	0.543	0.678	0.878	0.727
Desai et al., ICCV09	0.824	0.617	0.680	0.881	0.807
Our method	**0.871 ±0.018**	**0.793 ±0.012**	**0.719 ±0.014**	**0.897 ±0.021**	**0.820 ±0.011**

Average precision of object detection for the LabelMe dataset

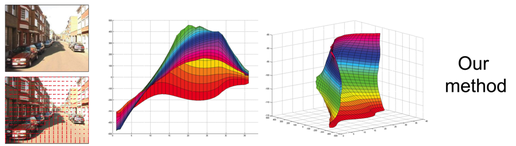


Zoomed-in detail

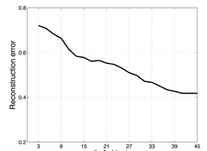


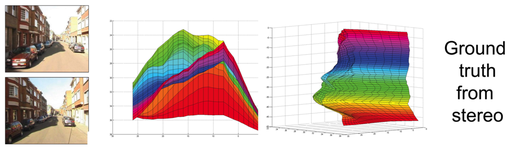
Detections | Normals | Discretized normals | Gupta et al. [3]

Comparison to [3] on the Geometric Context dataset.

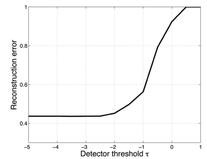


Our method





Ground truth from stereo



Scene reconstruction for the Leuven dataset.