Looking Back, Moving Forward

Computational Photography

Derek Hoiem, University of Illinois

Photo Credit
Lee Cullivan

Today

- Beyond this class…
- ICES forms
- Reminder: final project
  - Reports due Dec 14 11:59pm
  - Poster presentations on Dec 15 at 1:30pm on the first floor of Siebel
    - Half of class will present at one time, then switch
    - Everyone is assigned to review two posters (and should also look at the others that are of interest)

Project 5

- Incomplete list of excellent projects
  - http://tliang7.web.engr.illinois.edu/cs445/proj5/ -- Good smooth blend results (cut-car is removed)
  - http://geellis2.web.engr.illinois.edu/cs445/proj5/ -- Good additional video result
  - http://xwu68.web.engr.illinois.edu/cs445/proj5 -- Good additional video result
  - http://jmakdah2.web.engr.illinois.edu/cs445/proj5/ -- Good additional video result
  - http://kurtovc2.web.engr.illinois.edu/cs445/proj5/ -- Good explanation
  - http://dsun18.web.engr.illinois.edu/cs445/proj5/ -- Good results for background/foreground movie

This course has provided fundamentals

- How photographs are captured from and relate to the 3D scene

- How to think of an image as: a signal to be processed, a graph to be searched, an equation to be solved

- How to manipulate photographs: cutting, growing, compositing, morphing, stitching

- Basic principles of computer vision: filtering, correspondence, alignment

What else is out there?

Lots!

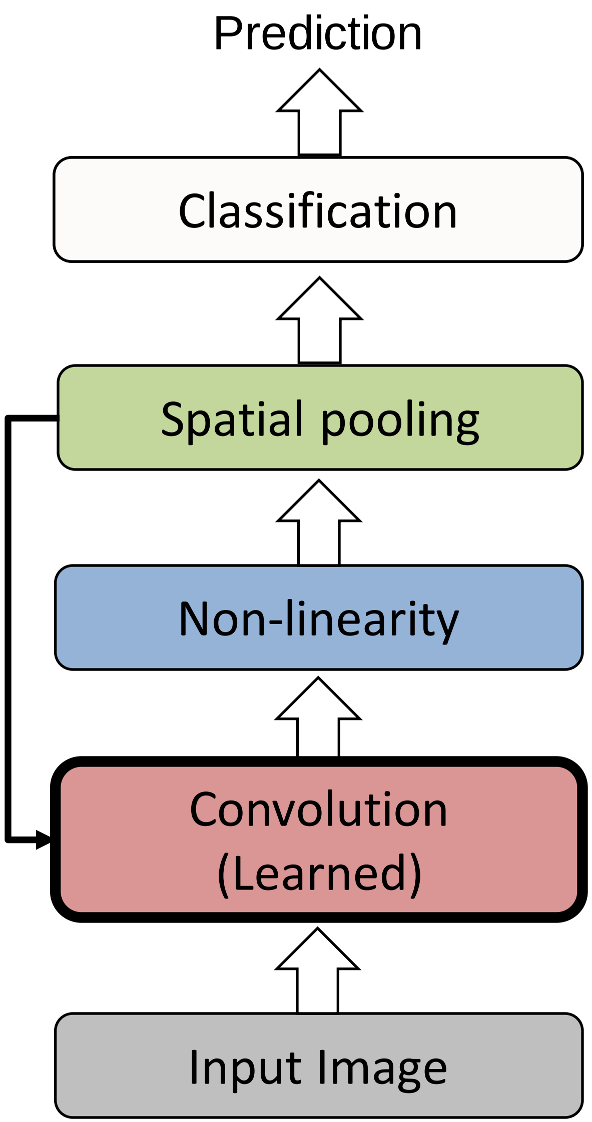
- Machine learning
- Videos and motion
- Scene understanding
- Modeling humans
- Better/cheaper devices
- …

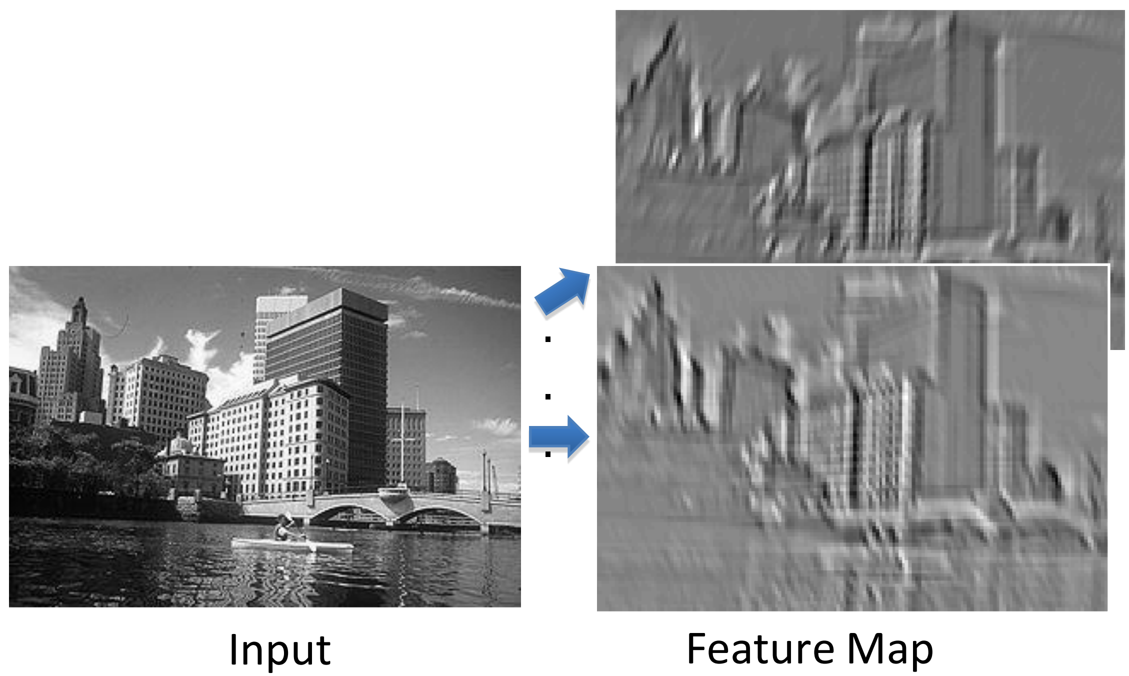
SIGGRAPH 2017 highlights

http://s2017.siggraph.org/technical-papers

Deep networks: new major influence

Key operations in a CNN

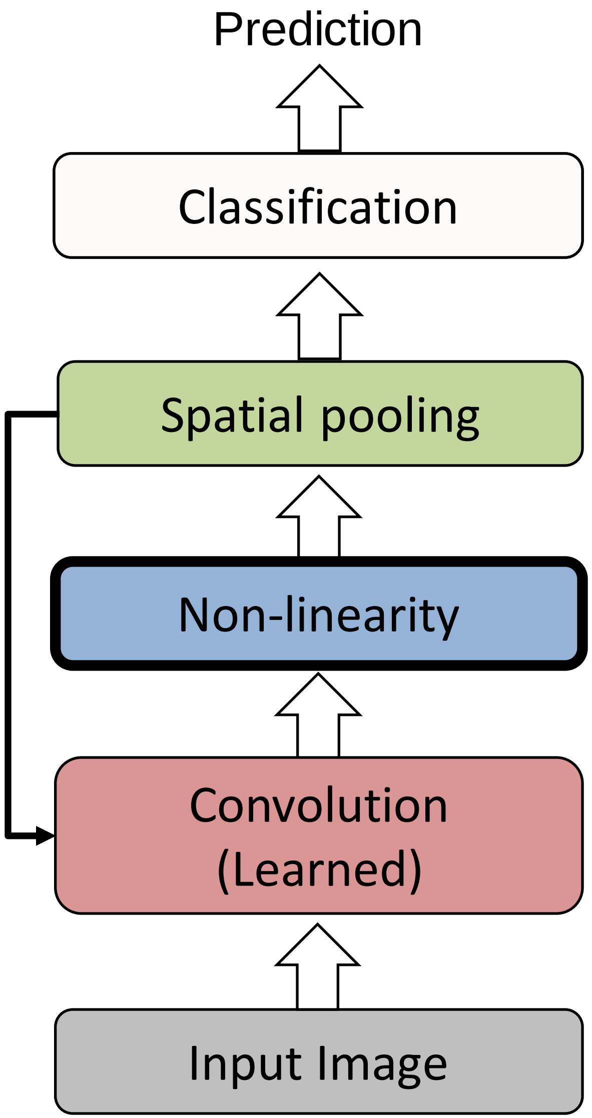




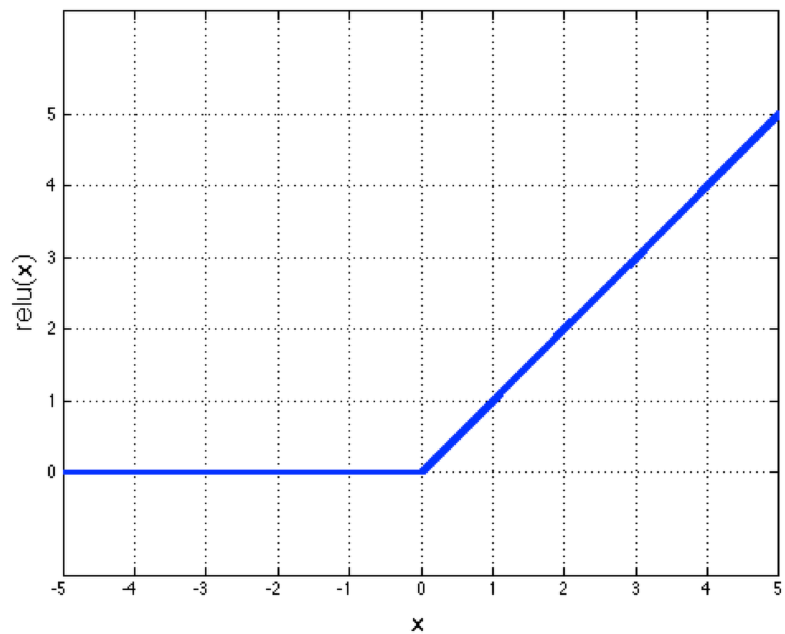
Source: R. Fergus, Y. LeCun

Slide: Lazebnik

Key operations



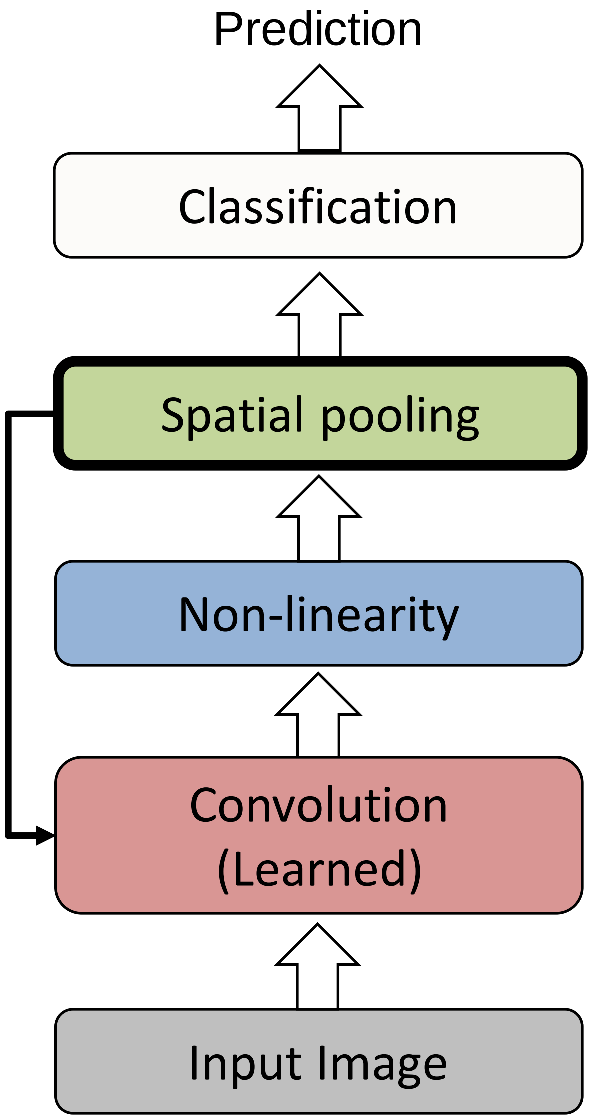
Rectified Linear Unit (ReLU)

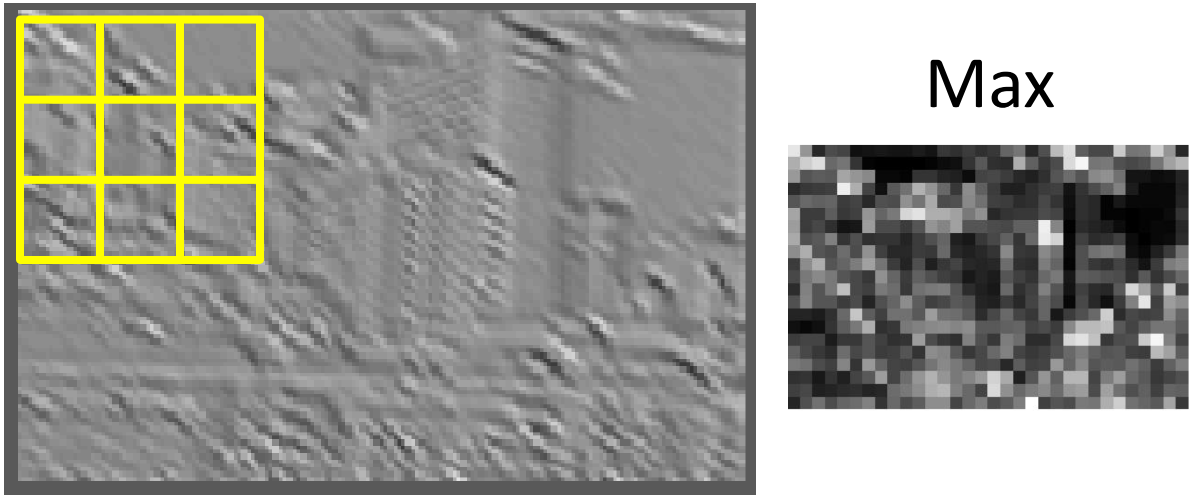


Source: R. Fergus, Y. LeCun

Slide: Lazebnik

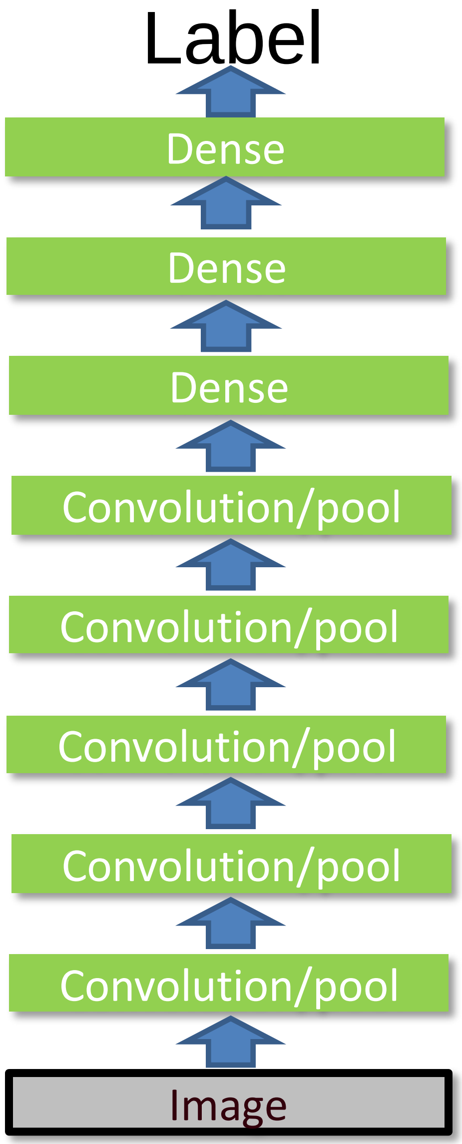
Key operations



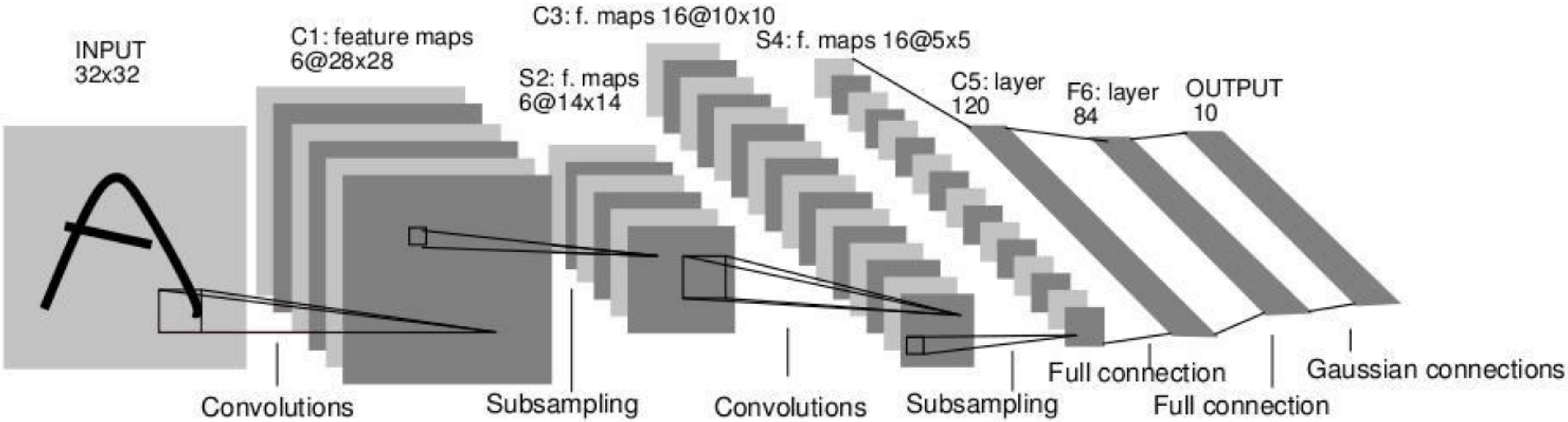


Source: R. Fergus, Y. LeCun

Key idea: learn features and classifier that work well together ("end-to-end training")



LeNet-5



- Average pooling
- Sigmoid or tanh nonlinearity
- Fully connected layers at the end
- Trained on MNIST digit dataset with 60K training examples

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86(11): 2278–2324, 1998.

Image-to-Image Translation with Conditional Adversarial Networks

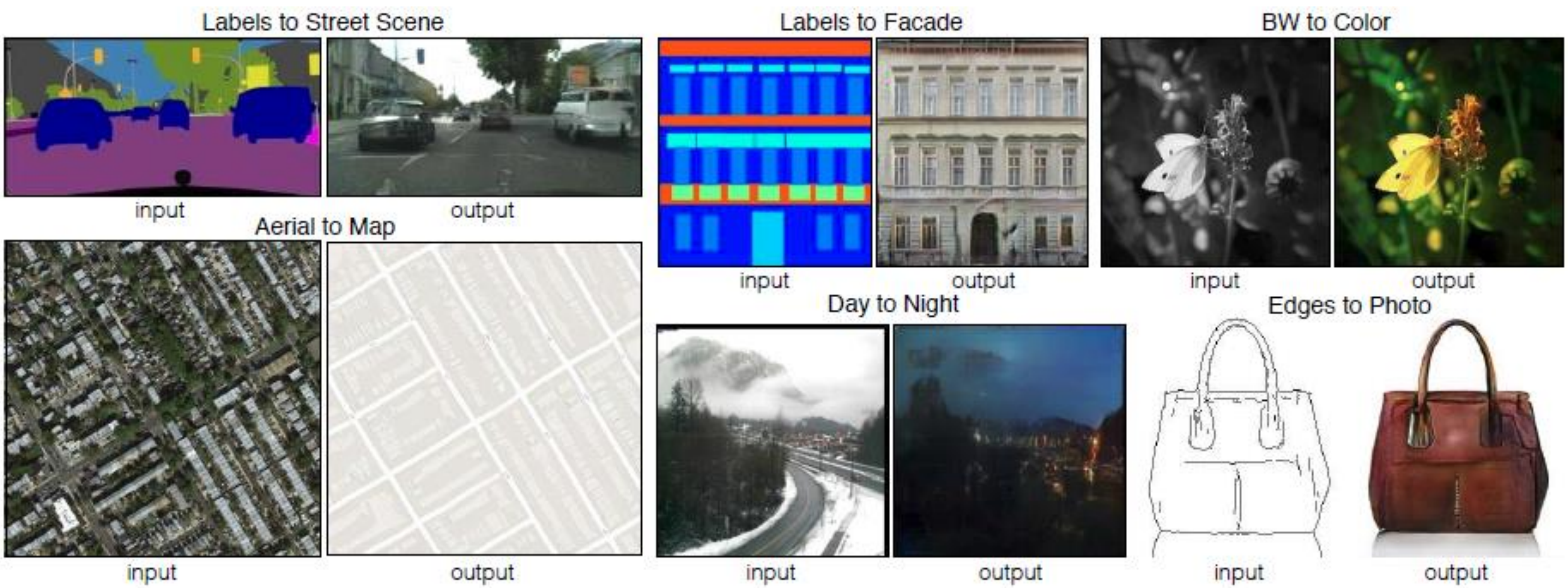
https://phillipi.github.io/pix2pix/

Phillip Isola Jun-Yan Zhu Tinghui Zhou Alexei A. Efros

Berkeley AI Research (BAIR) Laboratory
University of California, Berkeley
{isola,junyanz,tinghuiz,efros}@eecs.berkeley.edu

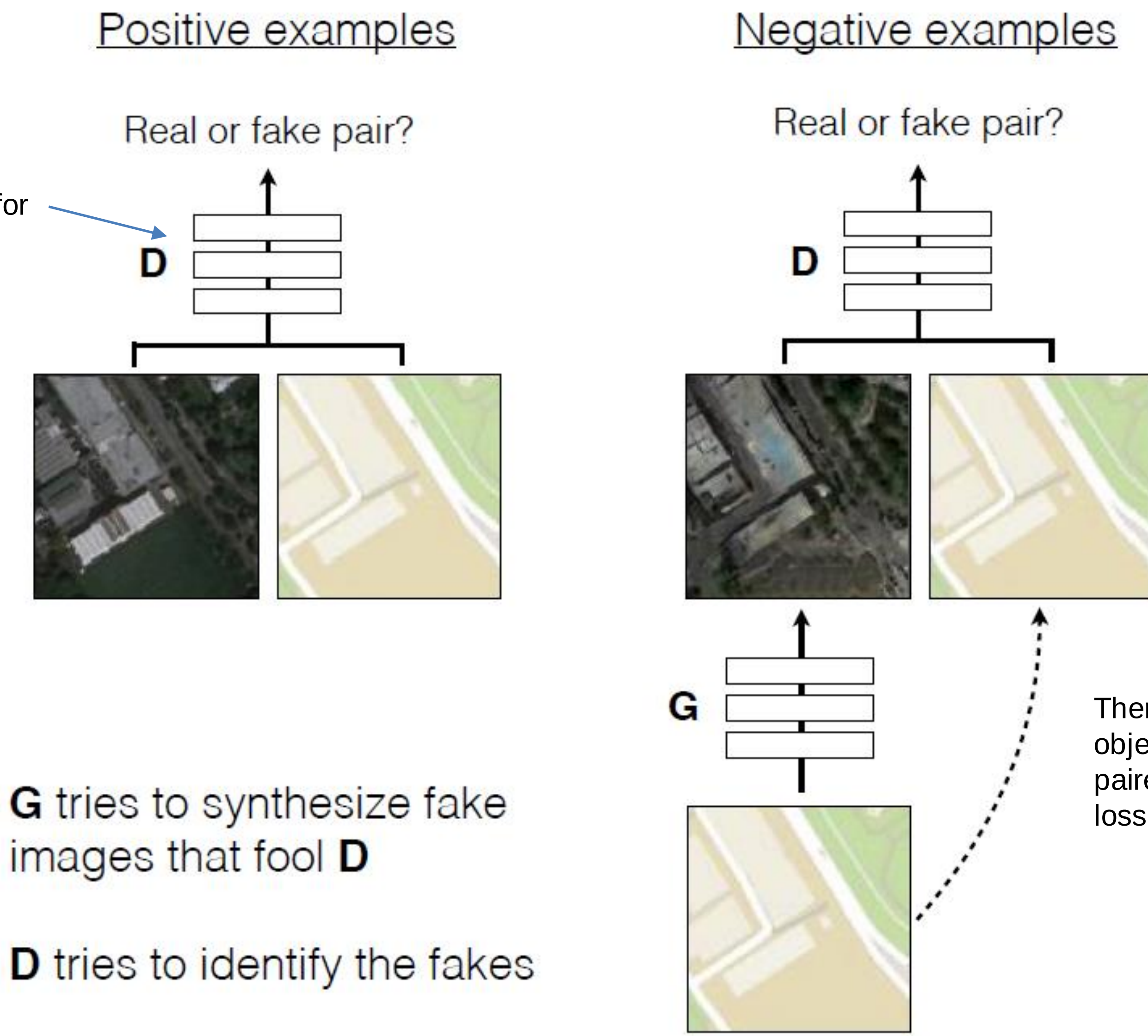
Learn to map from one image representation to another

- Trained from input/output pairs
- Patch memorization is implicit through learned representation



Learning to synthesize

Scores NxN patches for realism

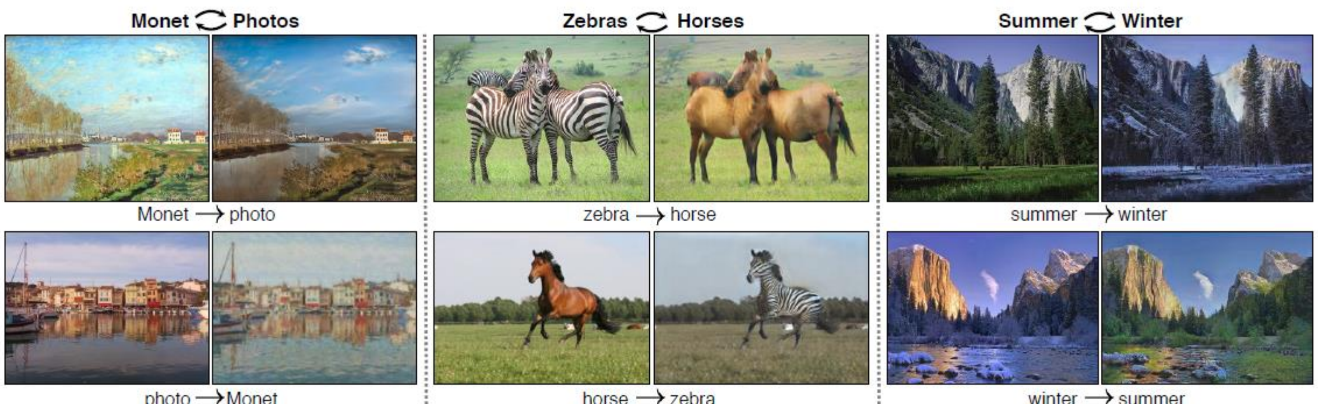


There is also an objective to produce the paired image with a L1 loss

CycleGAN

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Jun-Yan Zhu* Taesung Park* Phillip Isola Alexei A. Efros
Berkeley AI Research (BAIR) laboratory, UC Berkeley

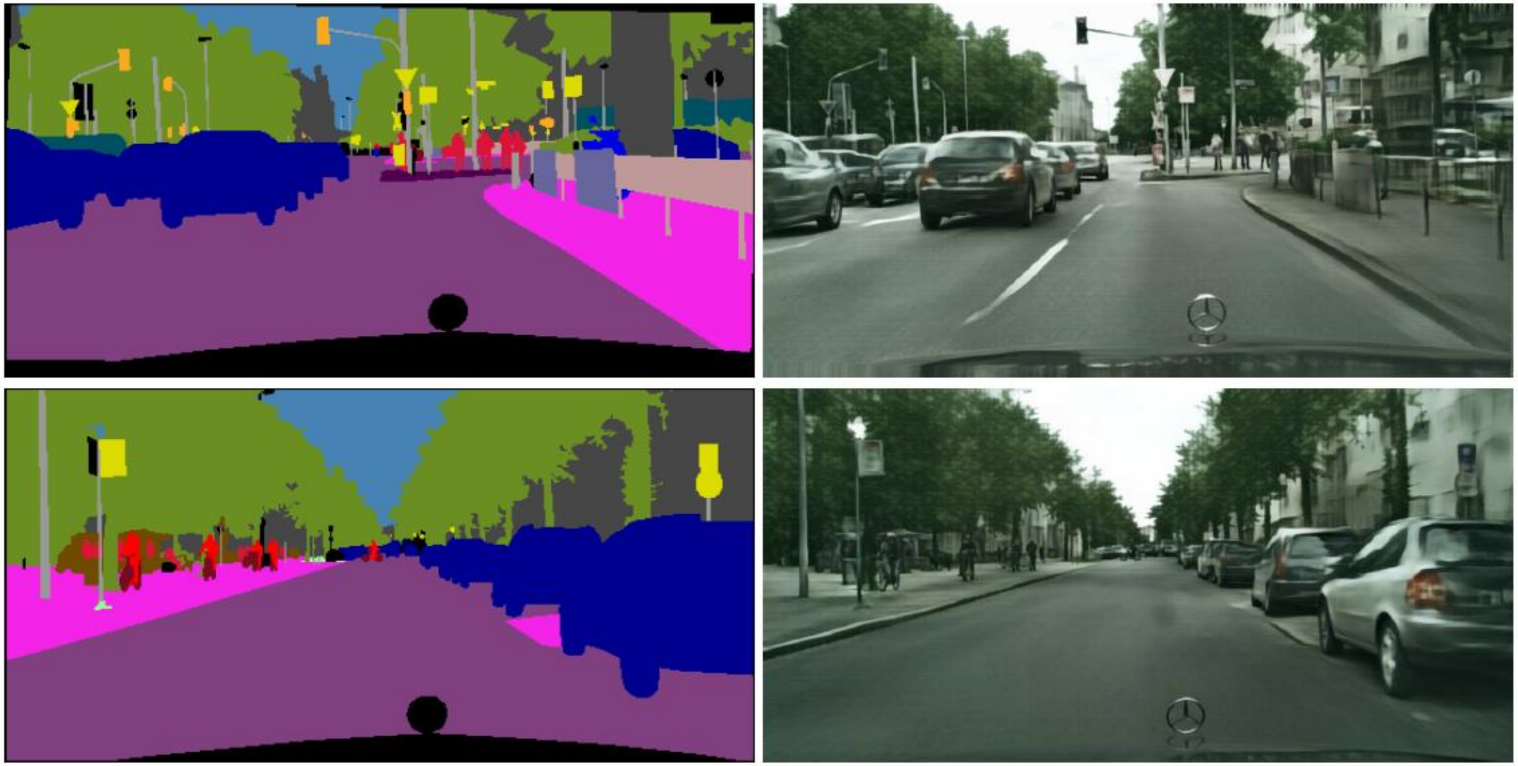


https://www.youtube.com/watch?v=yLCvWoQLnms

Image Synthesis

Photographic Image Synthesis with Cascaded Refinement Networks

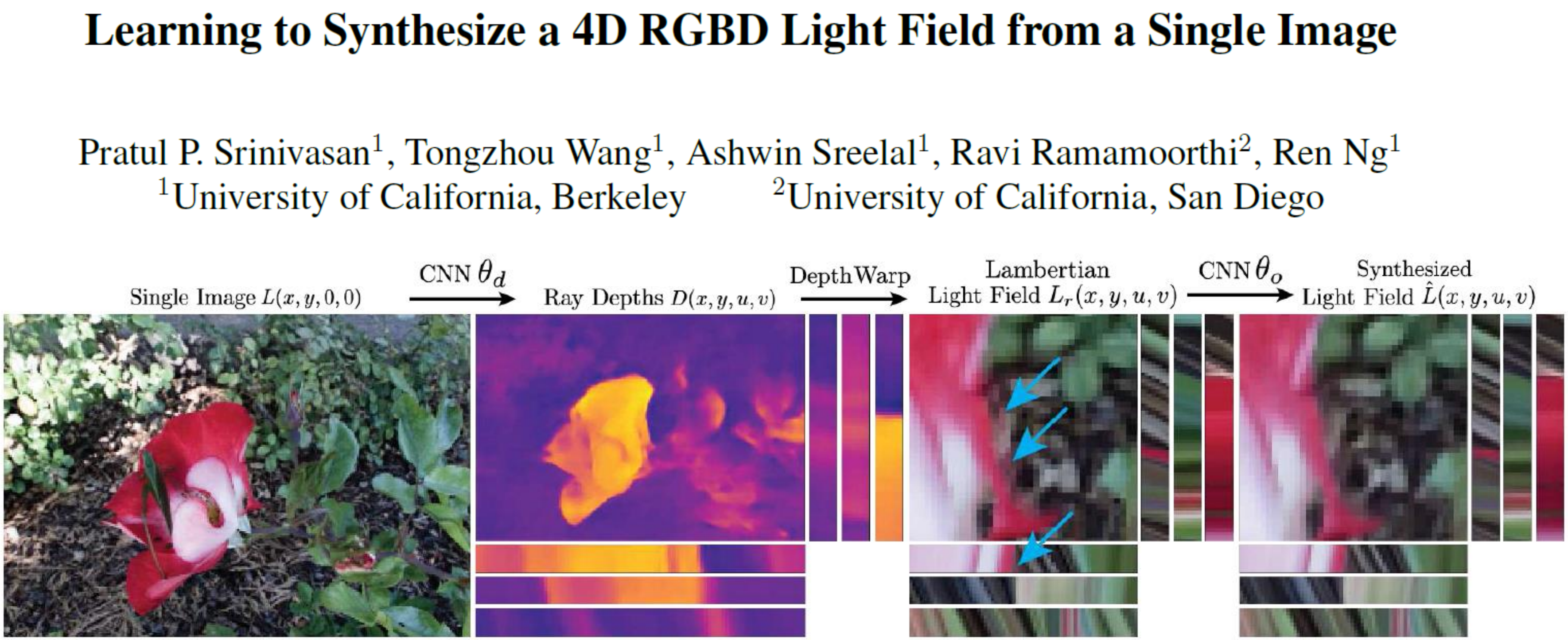
Qifeng Chen[† ‡] Vladlen Koltun[†]



(a) Input semantic layouts (b) Synthesized images

Cascaded Refinement Network (iteratively upsample features and refine, no GAN) + L1 loss on VGG features

Image → Light Field



https://www.youtube.com/watch?v=yLCvWoQLnms

Superresolution

EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis

Mehdi S. M. Sajjadi Bernhard Schölkopf Michael Hirsch

Bicubic ENet-E ENet-PAT Ground Truth

E: Optimize least squares objective with upsampling network
PAT: Optimize “perceptual” (VGG features) loss, adversarial loss, texture corr loss

(a) Input (b) SR [18] (c) SR [18]+Deblur [33] (d) Deblur [33] (e) Deblur [33]+SR [18] (f) Ours (g) GT

Learning to Super-Resolve Blurry Face and Text Images

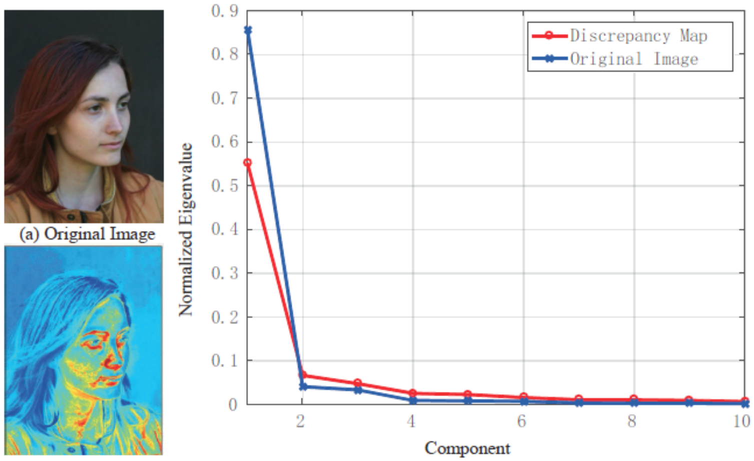
Xiangyu Xu[1,2,3] Deqing Sun[3,4] Jinshan Pan[5] Yujin Zhang[1]
Hanspeter Pfister[3] Ming-Hsuan Yang[2]
[1]Tsinghua University [2]University of California, Merced [3]Harvard University
[4]Nvidia [5]Nanjing University of Science & Technology

Pretty similar to above, more limited domain

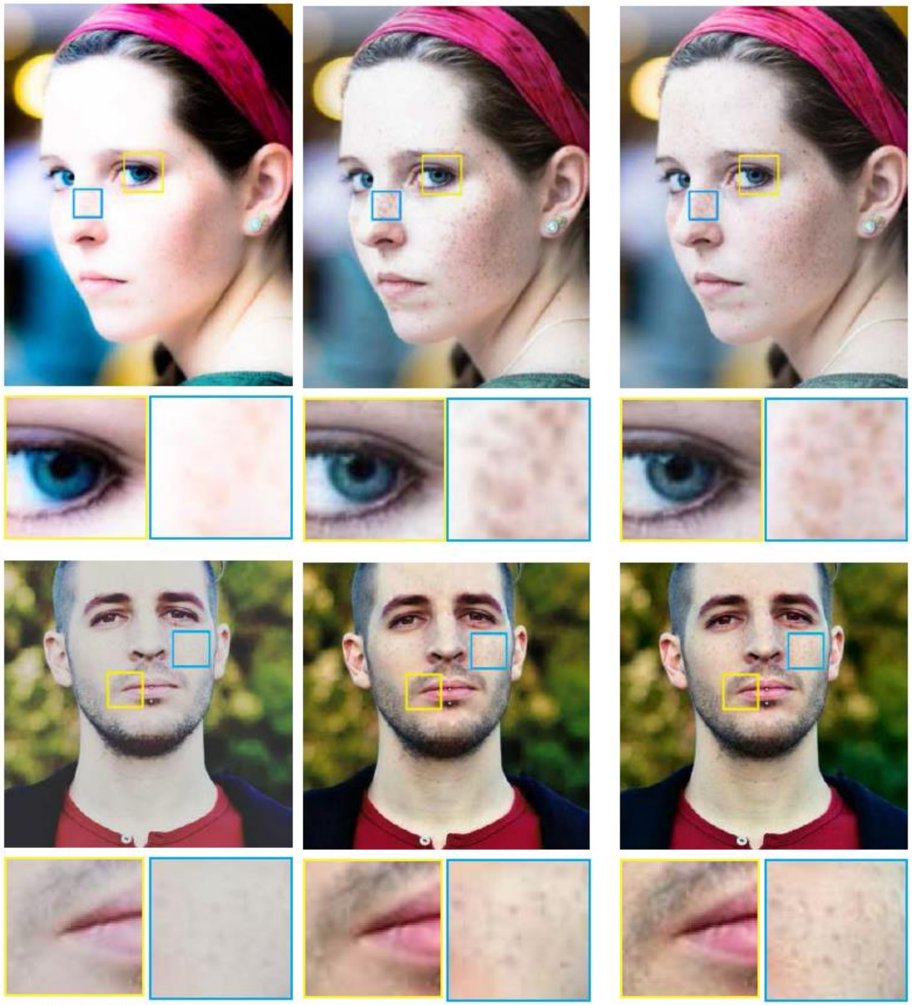
De-beautification

Makeup-Go: Blind Reversion of Portrait Edit*

Ying-Cong Chen[1] Xiaoyong Shen[2] Jiaya Jia[1,2]
[1]The Chinese University of Hong Kong [2]Tencent Youtu Lab
ycchen@cse.cuhk.edu.hk dylanshen@tencent.com leojia9@gmail.com



(a) Original Image (b) Discrepancy Map (c) Normalized Eigenvalue Distribution



(a) Edited Image (d) Ours (e) Ground truth

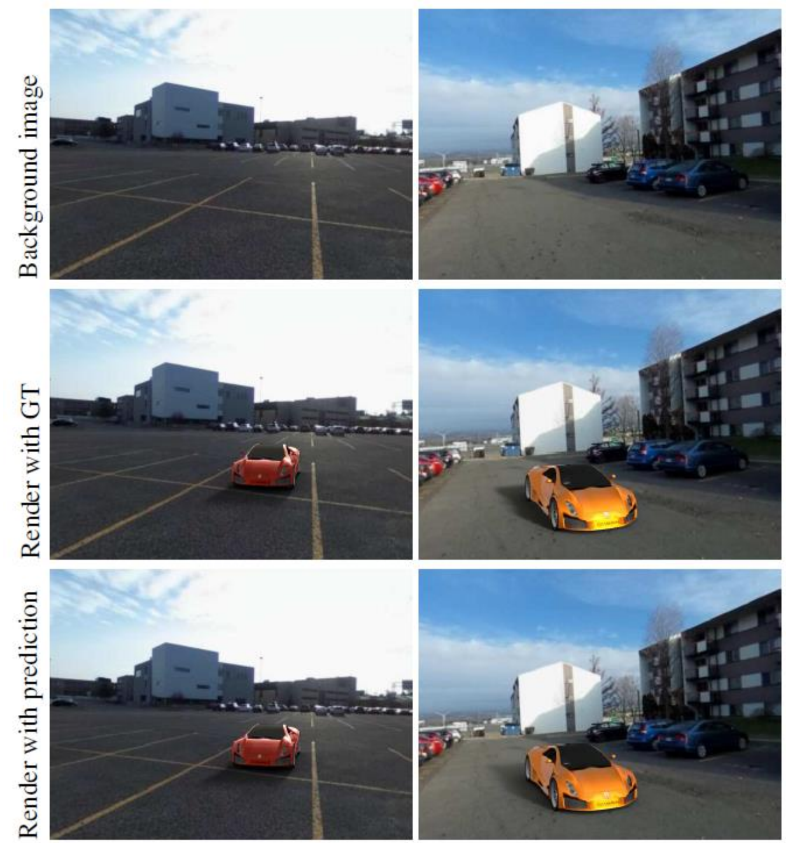
Network regresses principal components of discrepancy map

LDR --> HDR

Learning High Dynamic Range from Outdoor Panoramas

Jinsong Zhang Jean-François Lalonde
Université Laval, Québec, Canada
jinsong.zhang.1@ulaval.ca, jflalonde@gel.ulaval.ca
http://www.jflalonde.ca/projects/learningHDR

- Regress HDR from one LDR image
- Train on synthetic data
- Limited to outdoor scenes, rotated so that sun is on top



Smarter user assistance

- Handwriting beautification (Zitnick SG'13)
- 3D object modeling (Chen et al. SGA'13)
- 3D object modeling (Kholgade et al. SG'14)

Video and motion

- Video = sequence of images
  - Track points → optical flow, tracked objects, 3D reconstruction
  - Find coherent space-time regions → segmentation
  - Recognizing actions and events
- Examples:
  - Point tracking for structure-from-motion
    - Boujou 1
  - Facial transfer: Xu et al. SG2014

Scene understanding

Interpret image in terms of scene categories, objects, surfaces, interactions, goals, etc.

- Remove the guy lying down (Alyosha)
- Make the woman dance or the guy get up
- Fill in the window with bricks
- Find me images with only Alyosha and Piotro

Scene understanding

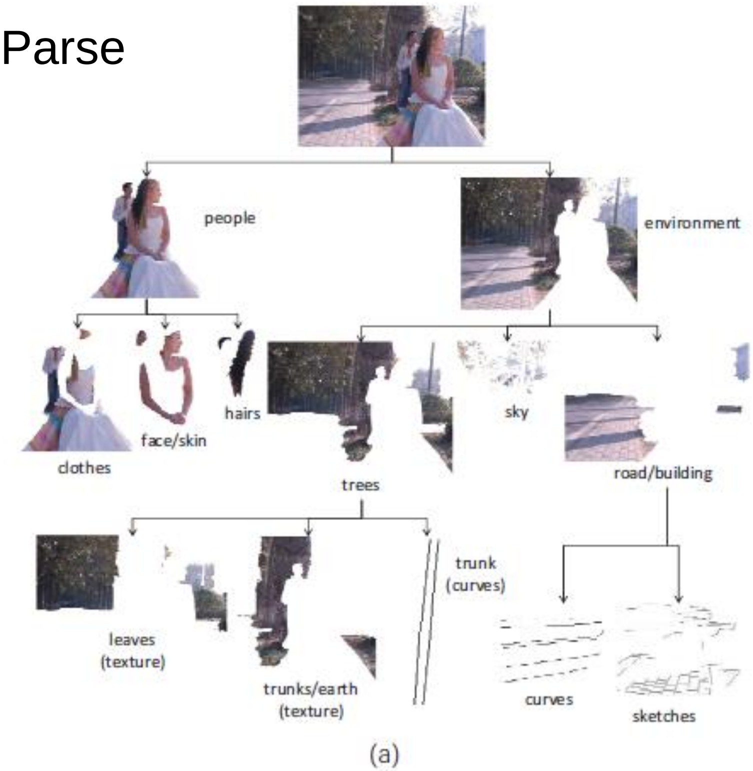
- Mostly unsolved, but we're getting there (especially for graphics purposes)
- Examples
  - "From Image Parsing to Painterly Rendering" (Zeng et al. 2010)
  - "Sketch2Photo: Internet Image Montage" (Chen et al. 2009)
  - Editing via scene attributes (Laffont et al. 2014)

Image Parsing to Painterly Rendering

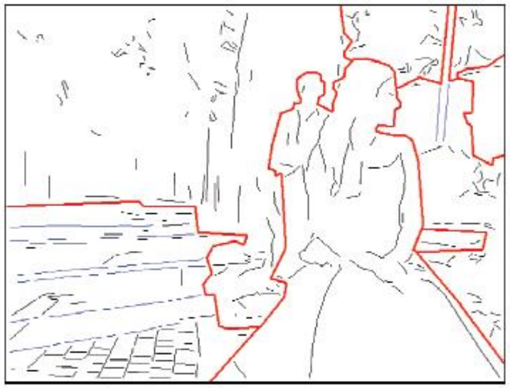
Zeng et al. SIGGRAPH 2010

Image Parsing to Painterly Rendering

Parse



Sketch



Brush Orientations

Brush Strokes

Zeng et al. SIGGRAPH 2010

Image Parsing to Painterly Rendering

Zeng et al. SIGGRAPH 2010

Image Parsing to Painterly Rendering

More examples

- Sketch2photo: http://www.youtube.com/watch?v=dW1Epl2LdFM

- Animating still photographs

Chen et al. 2009

Modeling humans

- Estimating pose and shape
  - http://clothingparsing.com/
  - Parselets (Dong et al., ICCV 2013)

- Motion capture

- 3D face from image (Kemelmacher ICCV'13)

Better and simpler 3D reconstruction

MobileFusion (2015): https://youtu.be/8M_-lSYqACo

Questions, Looking Forward

- How can we get computers to understand scenes (make predictions, describe them, etc.)?

- How can we design programs where semi-smart computers and people collaborate?

- What if we just capture and store the whole visual world (think StreetView)?

- How will photography change when depth cameras become standard?

How can you learn more?

- Relevant courses
  - Production graphics (CS 419)
  - Machine learning (CS 446 and others)
  - Computer vision (CS 543)
  - Optimization methods (w/ David Forsyth)
  - Parallel processing / GPU
  - HCI, data mining, NLP, robotics

Computer vision (with Prof Lazebnik Spring 2018)

Similar stuff to CP

- Camera models, filtering, single-view geometry, light and capture

New stuff

- Mid-level vision
  - Edge detection, clustering, segmentation
- Recognition
  - Image features and classifiers
  - Object category recognition
  - Action/activity recognition
- Videos
  - Tracking, optical flow
  - Structure from motion
- Multi-view geometry

How do you learn more?

Explore!

Thank you!

ICES forms and Feedback