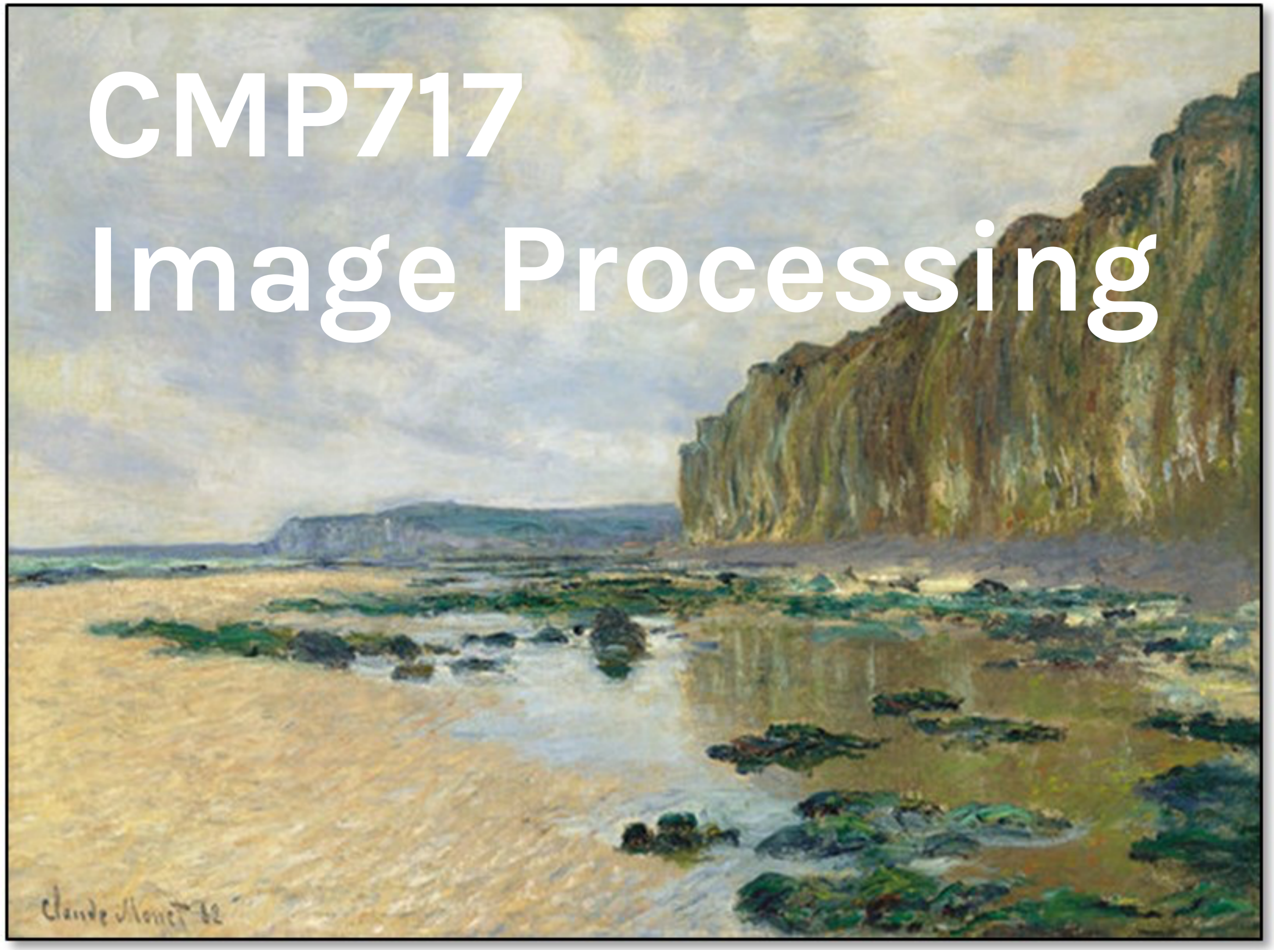
CMP717
Image Processing



Introduction

Erkut Erdem
Hacettepe University
Computer Vision Lab (HUCVL)

Today

- About me
- About you
- Course outline and logistics
- Introduction to Image Processing

Today

- About me
- About you
- Course outline and logistics
- Introduction to Image Processing

Erkut Erdem

Computer Vision Lab

http://web.cs.hacettepe.edu.tr/~erkut

My research interests concern computer vision and machine learning. I specifically investigate the use of spatial, temporal and cross-modal context in visual processing. My recent research activities cover topics such as saliency prediction, integrated vision and language, image editing and HDR image processing.

2001
BSc, Comp. Eng.
METU, Ankara

2003
MSc, Comp. Eng.
METU, Ankara

2004
Visiting Research Scholar
Virginia Tech, Blacksburg

2006
Visiting Student
UCLA, Los Angeles

2008
PhD, Comp. Eng.
METU, Ankara

2009
Post-doc
Telecom ParisTech (ENST), Paris

2012
Founder
Computer Vision Lab

2014
Asst. Prof.
Hacettepe, Ankara

2018
Assoc. Prof.
Hacettepe, Ankara

Structure Preserving Image Smoothing via Region Covariances

Levent Karacan,
Erkut Erdem,
Aykut Erdem

2013 ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2013)

Image Matting with KL-Divergence Based Sparse Sampling

Levent Karacan,
Aykut Erdem,
Erkut Erdem

2015 IEEE International Conference on Computer Vision (ICCV 2015)

Manipulating Attributes of Natural Scenes via Hallucination

Levent Karacan,
Zeynep Akata,
Aykut Erdem,
Erkut Erdem

2020 ACM Transactions on Graphics

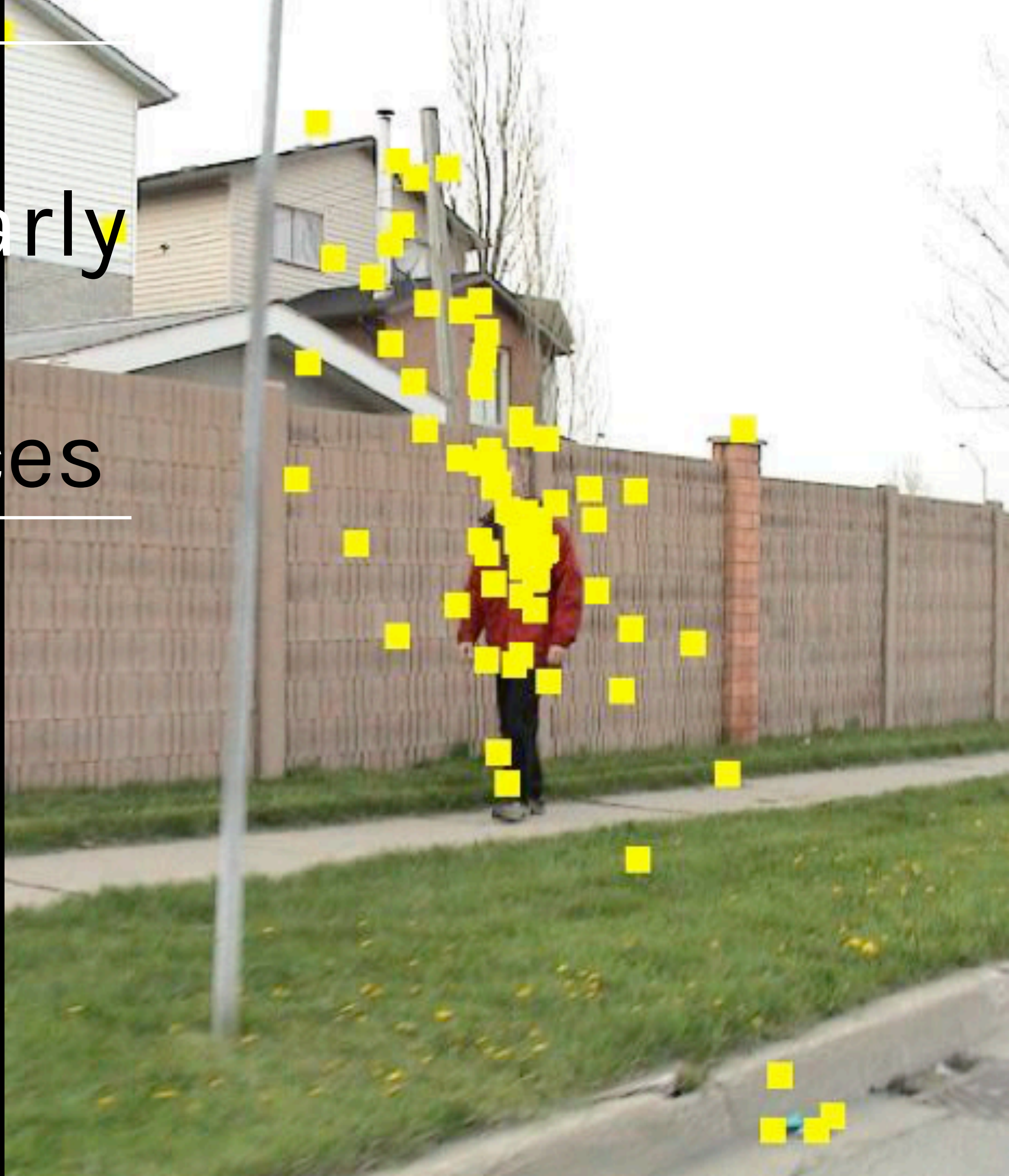
Alpha Matting with KL-Divergence Based Sparse Sampling

Levent Karacan,
Aykut Erdem,
Erkut Erdem

2017
IEEE Transactions on Image Processing

Visual saliency estimation by nonlinearly integrating features using region covariances

Erkut Erdem,
Aykut Erdem

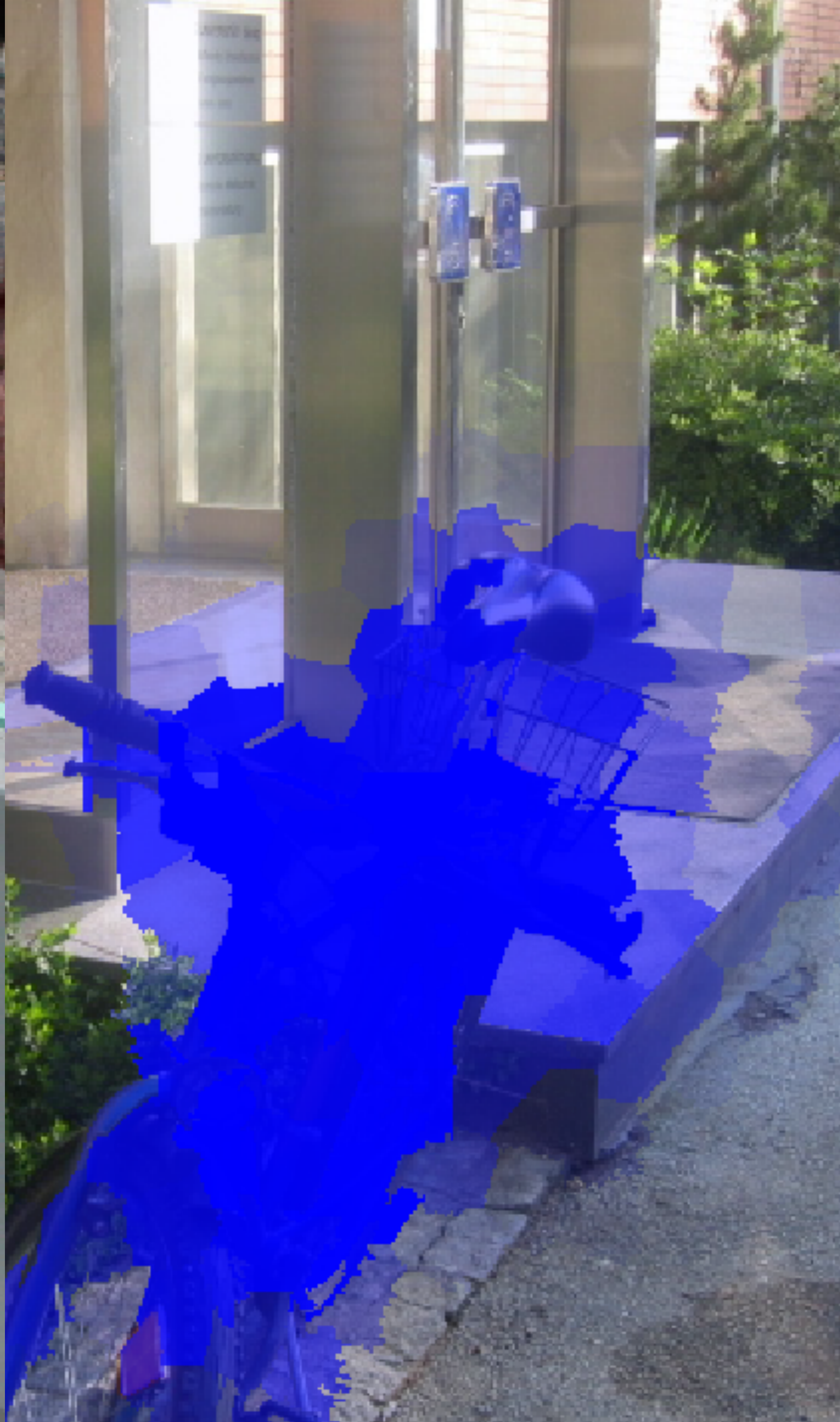


2013 Journal of Vision

Top down saliency estimation via superpixel based discriminative dictionaries

Aysun Kocak,
Kemal Cizmeciler,
Aykut Erdem,
Erkut Erdem

2014 British Machine Vision Conference (BMVC 2014)



A Comparative Study for Feature Integration Strategies in Dynamic Saliency Estimation

Yasin Kavak,
Erkut Erdem,
Aykut Erdem

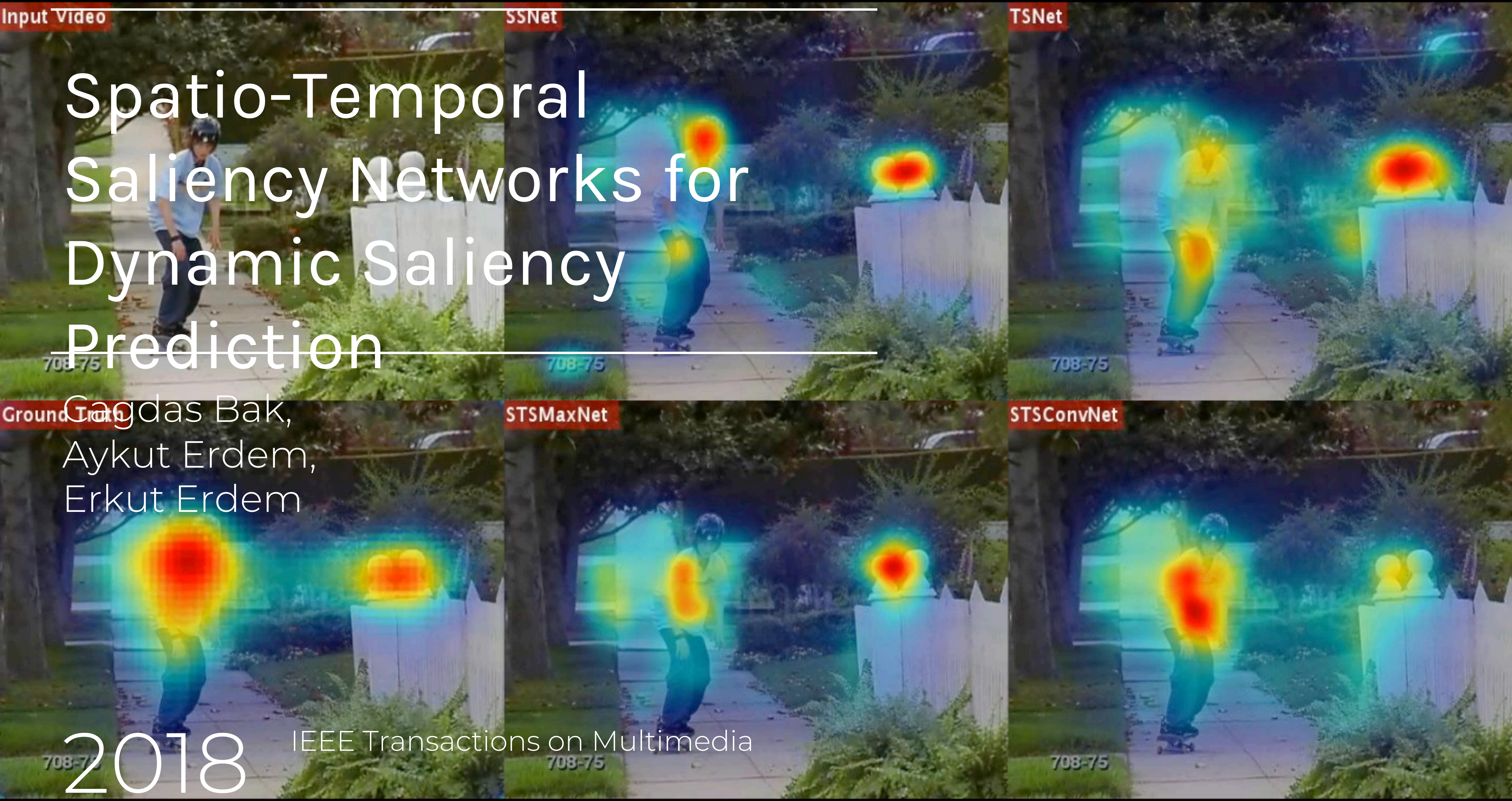
2017

Signal Processing: Image Communication

Spatio-Temporal Saliency Networks for Dynamic Saliency Prediction

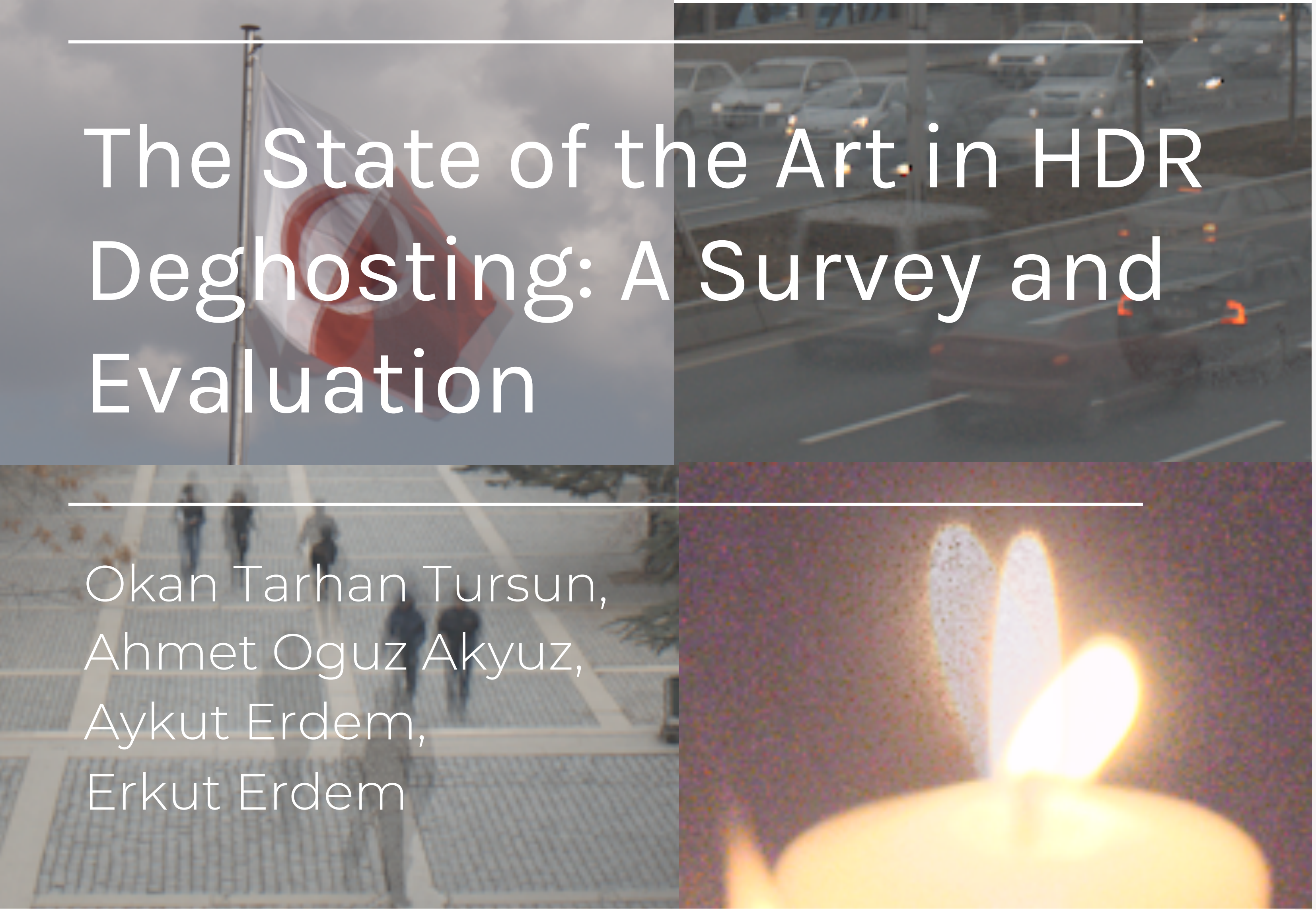
Cagdas Bak,
Aykut Erdem,
Erkut Erdem

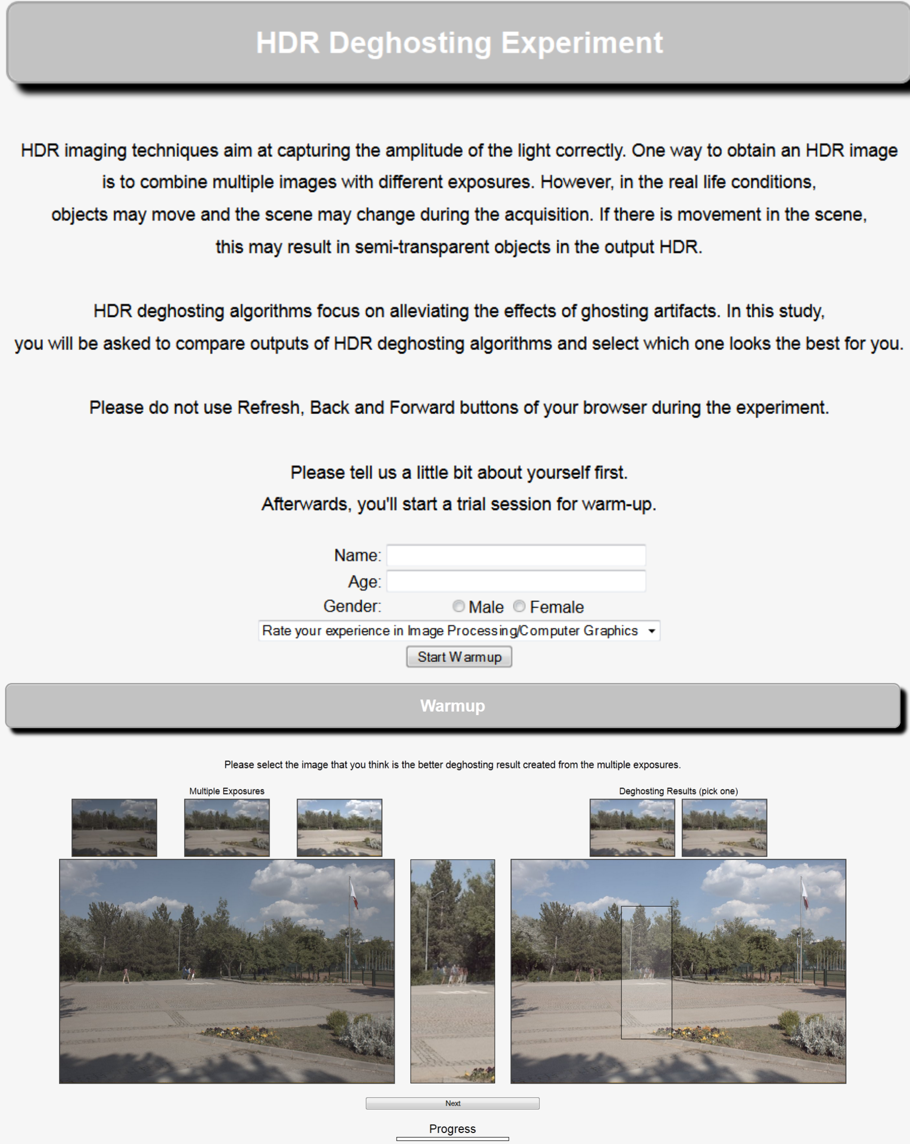
2018 IEEE Transactions on Multimedia



The State of the Art in HDR Deghosting: A Survey and Evaluation

Okan Tarhan Tursun,
Ahmet Oguz Akyuz,
Aykut Erdem,
Erkut Erdem





2015 Computer Graphics Forum (Eurographics STAR 2015)

Toplam Brut: 2.250,00 TL

Iyi calismalar.
Erkut

–

An Objective Deghosting Quality Metric for HDR Images

Okan Tarhan Tursun,
Ahmet Oguz Akyuz,
Aykut Erdem,
Erkut Erdem

2016 Computer Graphics Forum (Eurographics 2016)

Today

- About me
- About you
- Course outline and logistics
- Introduction to Image Processing

Introduce yourselves

- Who are you?
- What do you know about image processing?
- Why you want to take CMP717?

- Send me a short e-mail including your answers to these questions.

Today

- About me
- About you
- Course outline and logistics
- Introduction to Image Processing

Logistics

- Assoc. Prof. Erkut ERDEM
- erkut@cs.hacettepe.edu.tr
- Office: 112

- Lectures: Friday, 13:30-16:30
- Office Hour: *By appointment.*

About CMP717

- This course provides a comprehensive overview of fundamental topics in image processing for graduate students.
- The goal of this course is to provide a deeper understanding of the state-of-the-art methods in image processing literature and to study their connections.
- The course makes the students gain knowledge and skills in key topics and provides them the ability to employ them in their advanced-level studies.

Communication

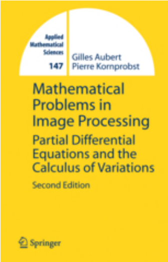
- The course webpage will be updated regularly throughout the semester with lecture notes, programming and reading assignments and important deadlines. http://web.cs.hacettepe.edu.tr/~erkut/cmp717.s20

- All other communications will be carried out through Piazza. Please enroll it by following the link https://piazza.com/hacettepe.edu.tr/spring2020/cmp717

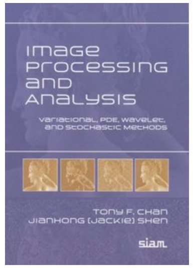
Prerequisites

- Good programming skills (for practicals and the course project)
- Calculus (differentiation, chain rule) and linear algebra (vectors, matrices, eigenvalues/vectors)
- Basic probability and statistics (random variables, expectations, multivariate Gaussians, Bayes rule, conditional probabilities)
- Undergraduate level image processing (e.g. BBM413)
- Machine learning (e.g. BBM406 and CMP712)
- Optimization (cost functions, taking gradients, regularization)

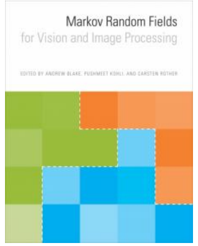
Reference Books



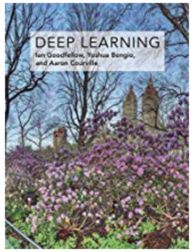
- Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, G. Aubert and P. Kornprobst, 2nd Edition, Springer-Verlag, 2006



- Image Processing And Analysis: Variational, PDE, Wavelet, And Stochastic Methods, T. Chan and J. Shen, Society for Industrial and Applied Mathematics, 2005



- Markov Random Fields For Vision And Image Processing, Edited by A. Blake, P. Kohli and C. Rother, MIT Press, 2011



- Deep Learning, Ian Goodfellow, Aaron Courville, and Yoshua Bengio, preparation for MIT Press,

Reading Material

- Lecture notes and handouts
- Papers and journal articles

Grading Policy

- Midterm Exam (24%)
- Paper Presentations (14%)
  (6% overview, 4% pros, and 4% cons)
- Weekly Quizzes (12%)
- Practical (8%)
- Course Project (presentations and reports) (42%)

Paper presentations and Quizzes

- An important part of the course includes discussions of a number papers related to certain research topics.
- These papers should be read by every student as the quizzes about the presented papers will be given on the weeks of the presentations.
- <u>The schedule for the presentations will be determined shortly.</u>

Structure of paper presentations

- Each paper discussion will be led by three students:
  - One student will be responsible from providing an overview of the paper.
  - One student will present the strengths of the paper.
  - One student will discuss the weaknesses of the paper.

Grading Rubric - Paper Overview

Criterion	Max	Points
Problem statement and motivation Clear definition of the problem, why it is interesting and important	10	
High-level overview of the paper Main contributions	10	
Key technical ideas Overview of the approach, related work	30	
Experimental set-up Datasets, evaluation metrics, applications	15	
Overall effectiveness of slide text/visuals Good balance of text and figures	10	
Overall effectiveness of the presentation Good oral skills, ability to answer follow-on questions, good leading of the class discussions	15	
Time Effective usage of time (~12 minutes long)	10	

Grading Rubric - Paper Strengths

Criterion	Max	Points
Summary of the paper One slide summary of the proposed approach	5	
Connections with other work How the method relates to other approaches	10	
Strengths of the approach Discuss the novelty of the approach, how it improves the existing work	25	
Strengths of the evaluation protocol Discuss the baselines and the ablation procedure	25	
Overall effectiveness of slide text/visuals Good balance of text and figures	10	
Overall effectiveness of the presentation Good oral skills, ability to answer follow-on questions, good leading of the class discussions	15	
Time Effective usage of time (~9 minutes long)	10	

Grading Rubric - Paper Weaknesses

Criterion	Max	Points
Summary of the paper One slide summary of the proposed approach	5	
Weaknesses of the approach Describe some cases in which you expect the approach performs poorly	25	
Weaknesses of the evaluation protocol Describe how the evaluation could be improved	25	
Future direction Open research questions, possible improvements over the approach	10	
Overall effectiveness of slide text/visuals Good balance of text and figures	10	
Overall effectiveness of the presentation Good oral skills, ability to answer follow-on questions, good leading of the class discussions	15	
Time Effective usage of time (~9 minutes long)	10	

Practical

- A single programming assignment that involves implementation, analysis, and reporting.
- Should be done individually
- 8% of your overall grade
- No late policy
- Out: April 3rd, Due: April 17th (these dates are tentative)

Project

- Aim: To give the students some experience on conducting research.
- Students should work individually or groups in two.
- This project may involve
  - Design of a novel approach and its experimental analysis,
  - an extension to a recent study of non-trivial complexity and its experimental analysis
- https://web.cs.hacettepe.edu.tr/~erkut/cmp717.s20/project.html

Project

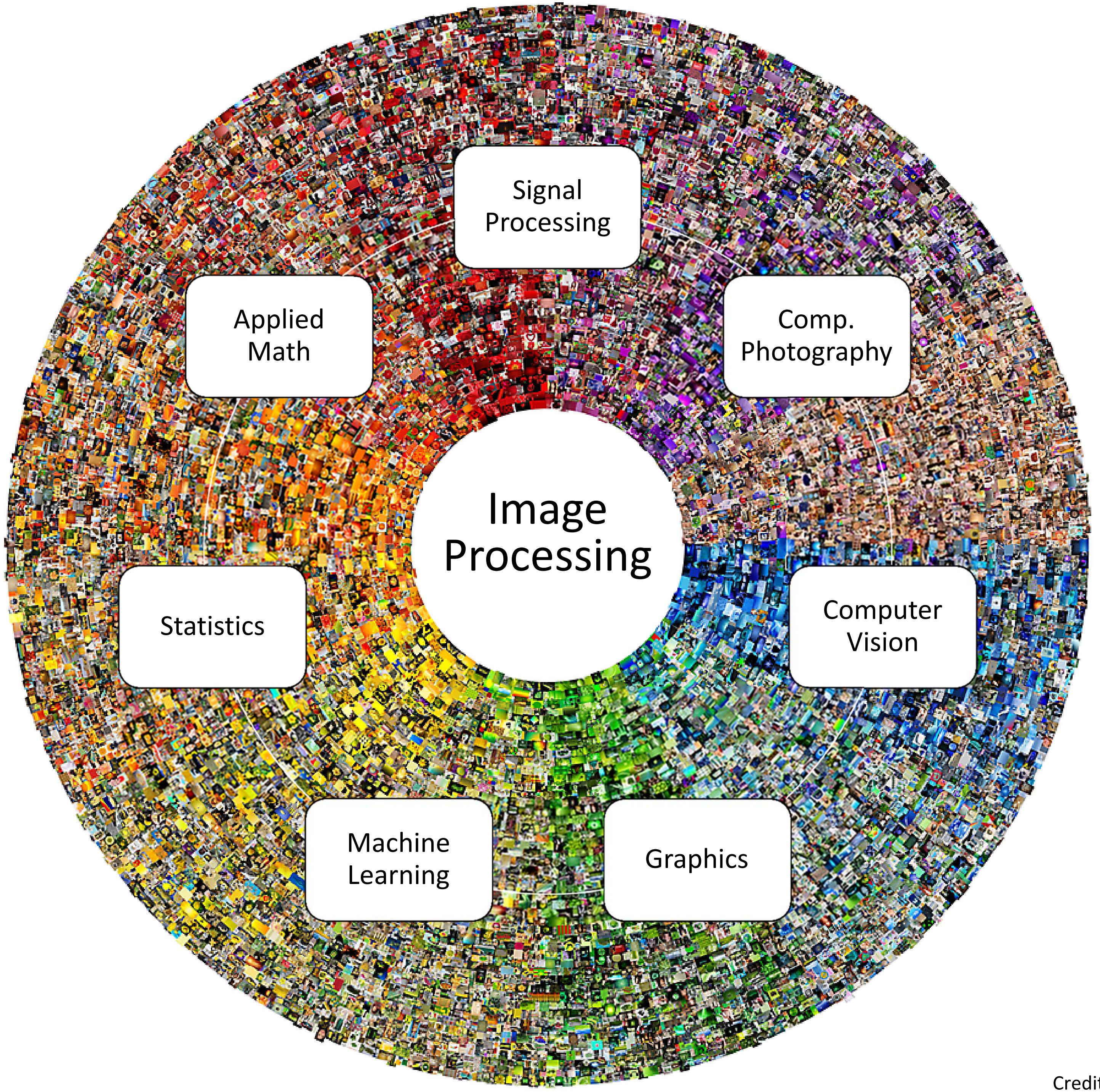
- Deliverables
  - Proposals: March 27
  - Project progress presentations: April 24
  - Project progress reports: May 1
  - Final project presentations: May 29
  - Final reports: June 15
- Grading
  - Proposal (4%)
  - Progress report (8%)
  - Progress presentation (8%)
  - Project presentation (10%)
  - Final report and code (12%)

Tentative Outline

- Overview of Image Processing
- Linear Filtering, Edge/Boundary Detection, Image Segmentation
- Nonlinear Filtering
- Sparse Coding
- Deep Learning Basics
- Convolutional Neural Networks
- Deep Generative Networks
- Image to Image Translation
- Image Deblurring
- Semantic Segmentation
- Visual Saliency

Today

- About me
- About you
- Course outline and logistics
- Introduction to Image Processing



Credit: P. Milanfar

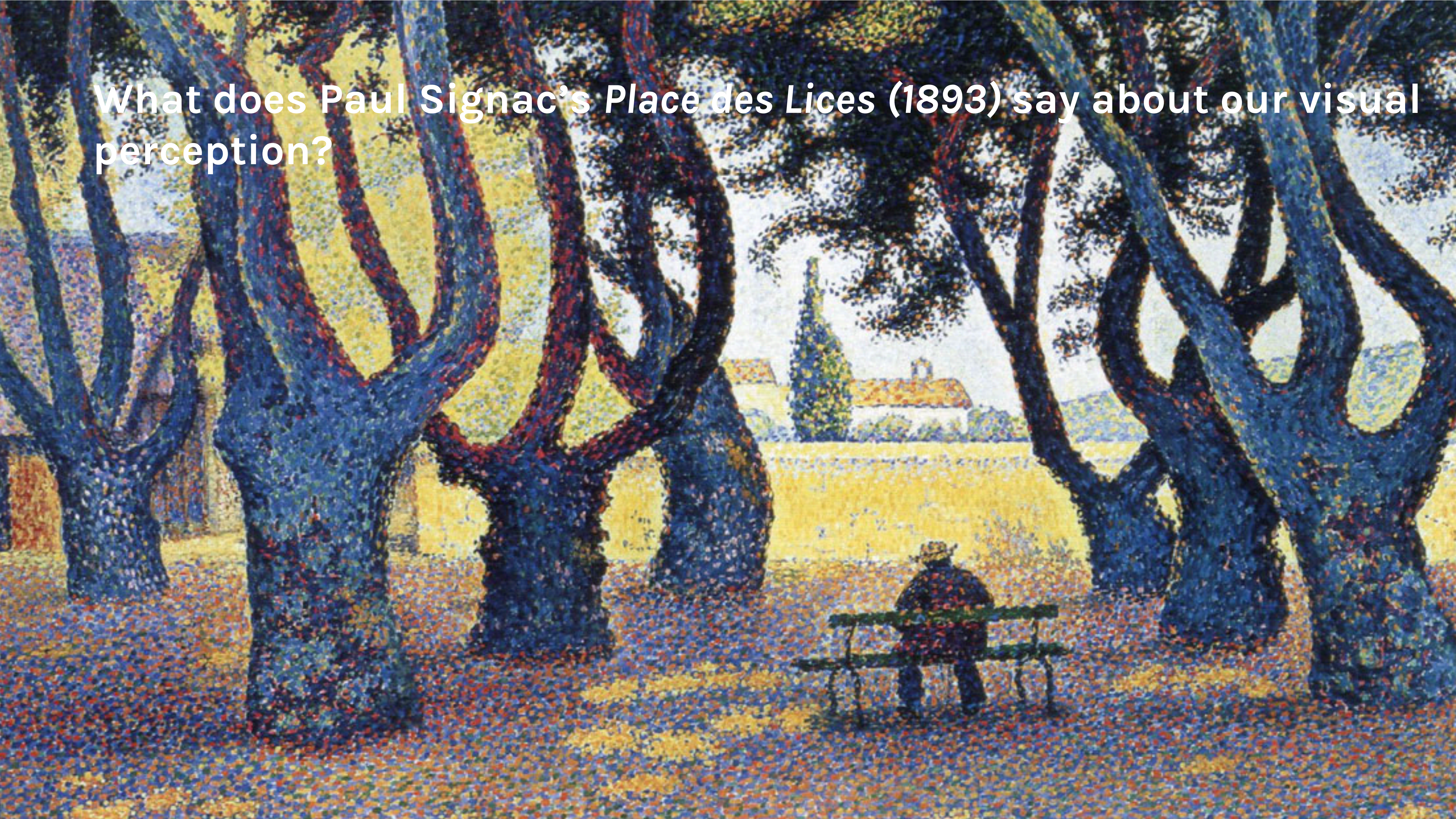
What does it mean, to see?

- *"The plain man's answer (and Aristotle's, too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is."* David Marr, Vision, 1982

- Our brain is able to use an image as an input, and interpret it in terms of objects and scene structures.

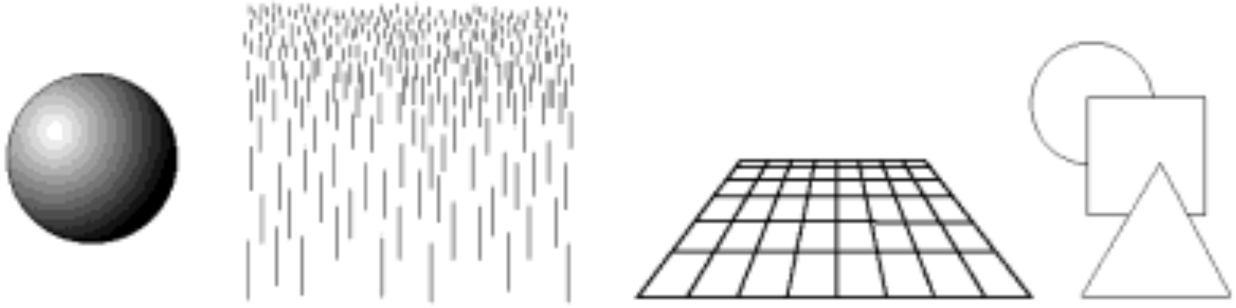
What does Salvador Dali's *Study for the Dream Sequence in Spellbound (1945)* say about our visual perception?

What does Paul Signac's *Place des Lices* (1893) say about our visual perception?



Why does vision appear easy to humans?

- Our brains are specialized to do vision.
- Nearly half of the cortex in a human brain is devoted to doing vision (cf. motor control ~20-30%, language ~10-20%)
- *"Vision has evolved to convert the <u>ill-posed problems</u> into solvable ones by adding premises: <u>assumptions</u> about how the world we evolved in is, on average, put together"*
  Steven Pinker, How the Mind Works, 1997
- Gestalt Theory (Laws of Visual Perception),
  Max Wertheimer, 1912



Figures: Steven Pinker, How the Mind Works, 1997

Computer Vision

- “Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information” ~David Marr

- <u>The goal of Computer Vision:</u>
  To develop artificial machine vision systems that make inferences related to the scene being viewed through the images acquired with digital cameras.

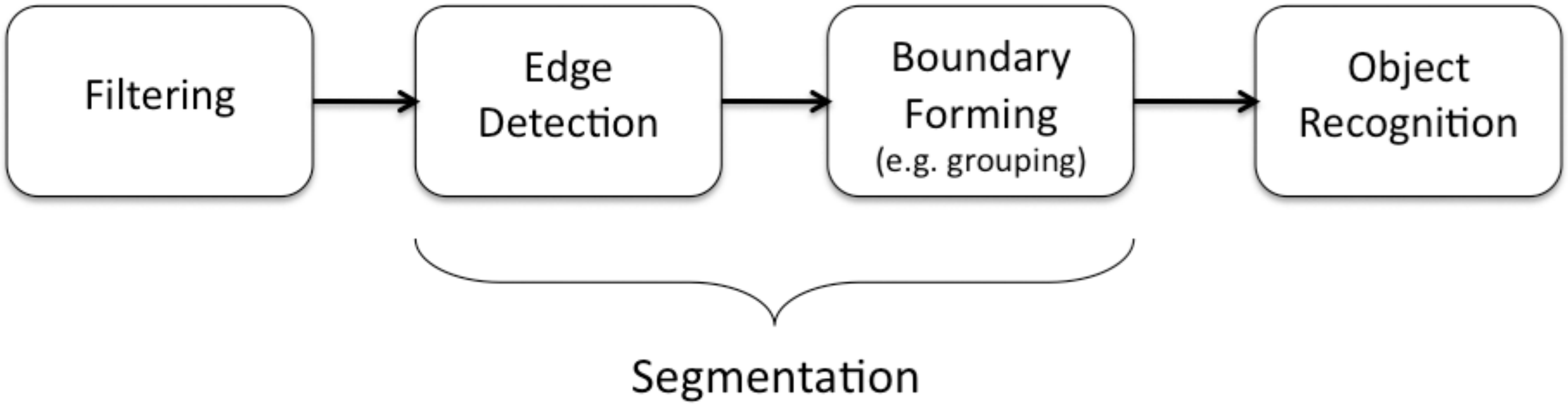
Marr's observation: Studying vision at three different levels

- Vision as an information processing task
  [David Marr, 1982]

- Three levels of understanding:

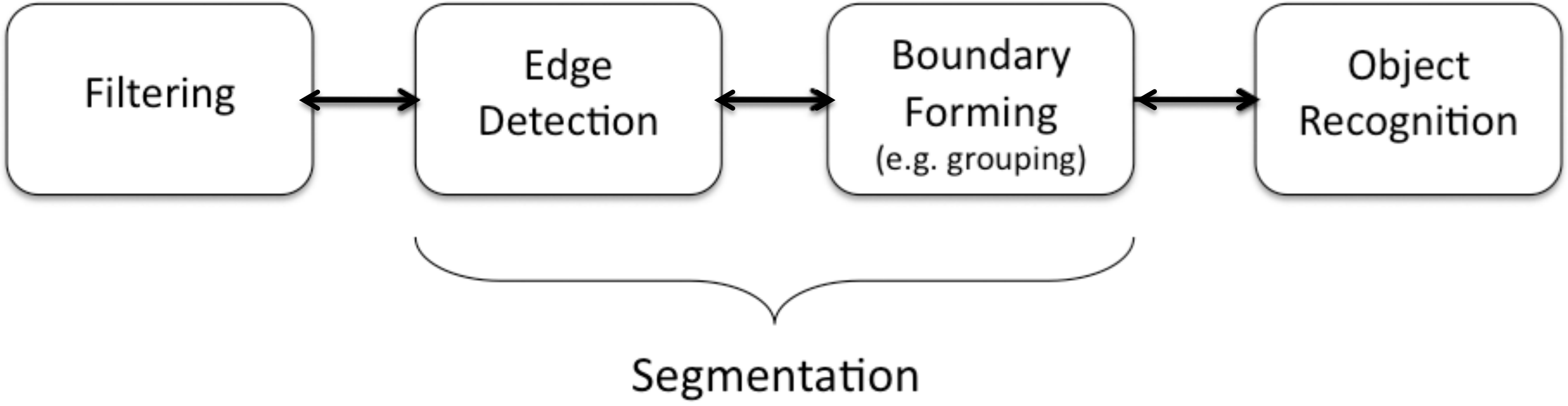
1. Computational theory
   - What is computed? Why it is computed?
2. Representation and Algorithm
   - How it is computed?
   - Input, Output, Transformation
3. Physical Realization
   - Hardware

Visual Modules and the Information Flow



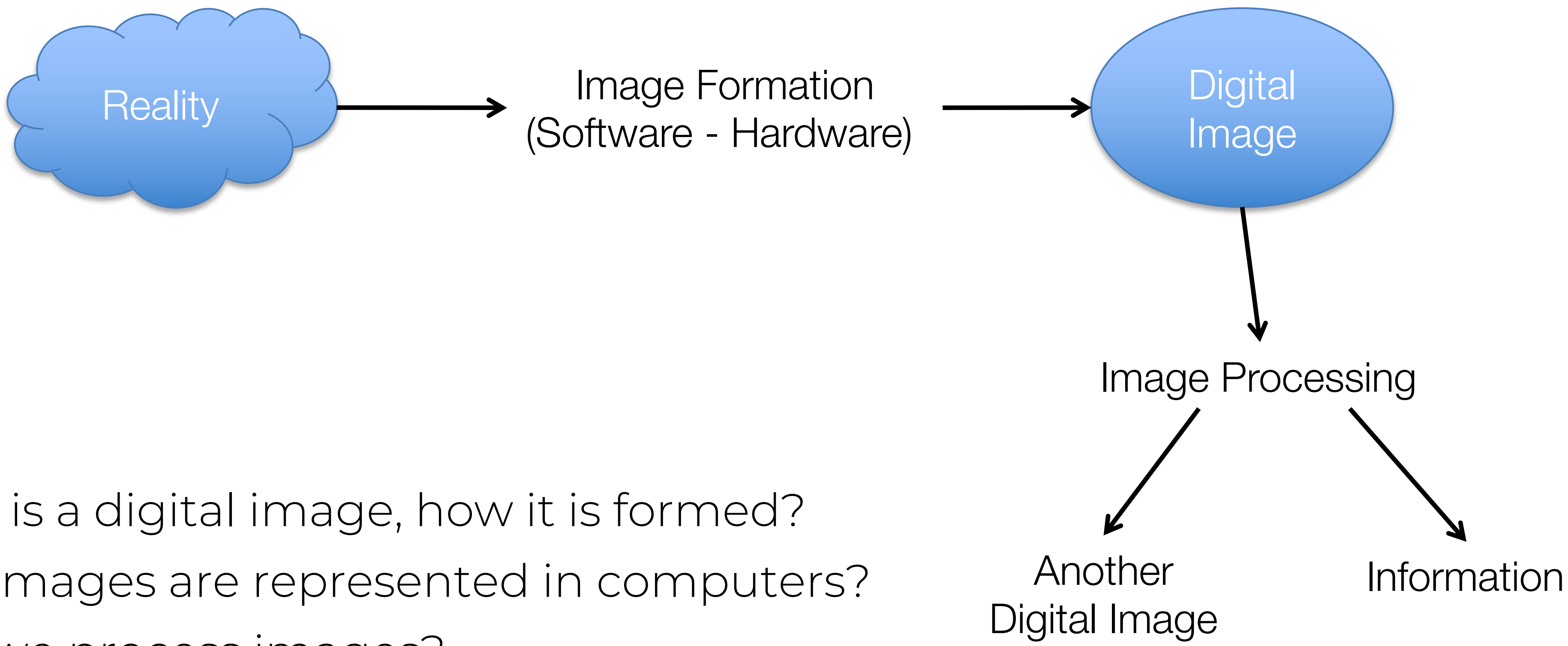
- Visual perception as a data-driven, bottom-up process (traditional view since D. Marr)
- Unidirectional information flow
- Simple low-level cues >> Complex abstract perceptual units

Visual Modules and the Information Flow



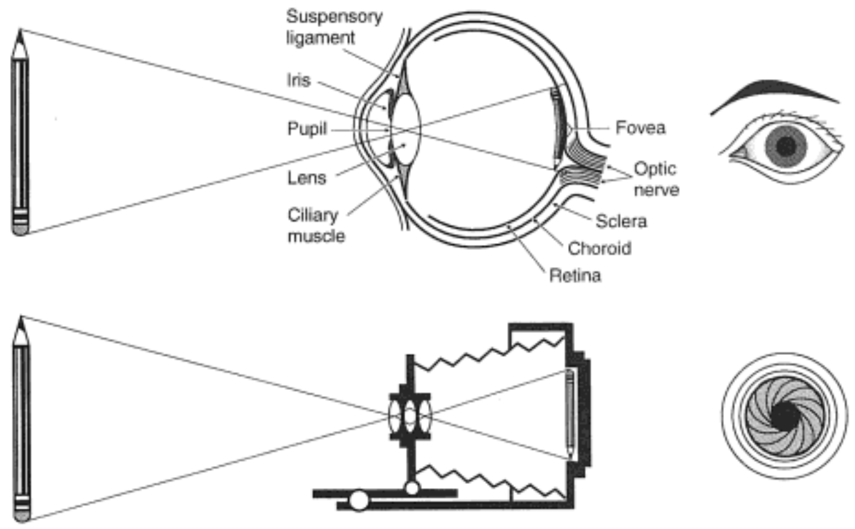
- Vision modules can be categorized into three groups according to their functionality:
  - Low-level vision: filtering out irrelevant image data
  - Mid-level vision: grouping pixels or boundary fragments together
  - High-level vision: complex cognitive processes

Fundamentals of Image Processing



- What is a digital image, how it is formed?
- How images are represented in computers?
- Why we process images?
- How we process images?

Image Formation



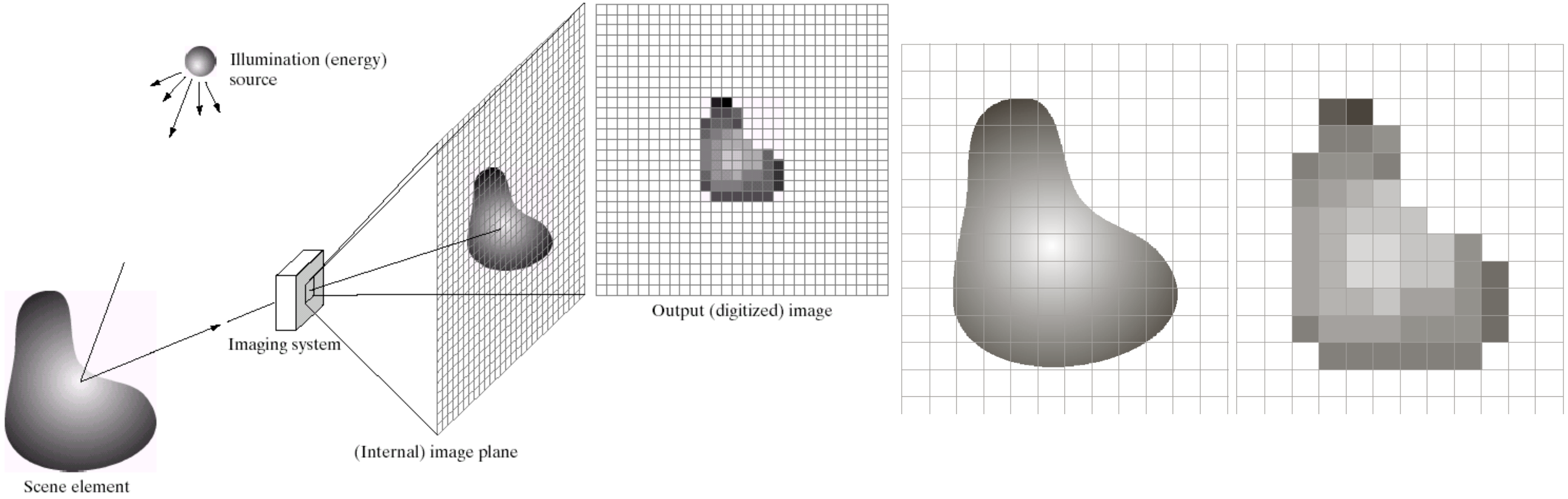
Three Dimensional World → Two Dimensional Image Space

- What is measured in an image location?
  - brightness
  - color

<< viewpoint
illumination conditions
local geometry
local material properties

Figures: Francis Crick, The Astonishing Hypothesis, 1995

Image Formation



- Discretization
  - in image space - sampling
  - In image brightness - quantization

Figures: Gonzalez and Woods, Digital Image Processing, 3rd Edition, 2008

Image Representation

- **Digital image:** 2D discrete function f
- **Pixel:** Smallest element of an image f(x,y)

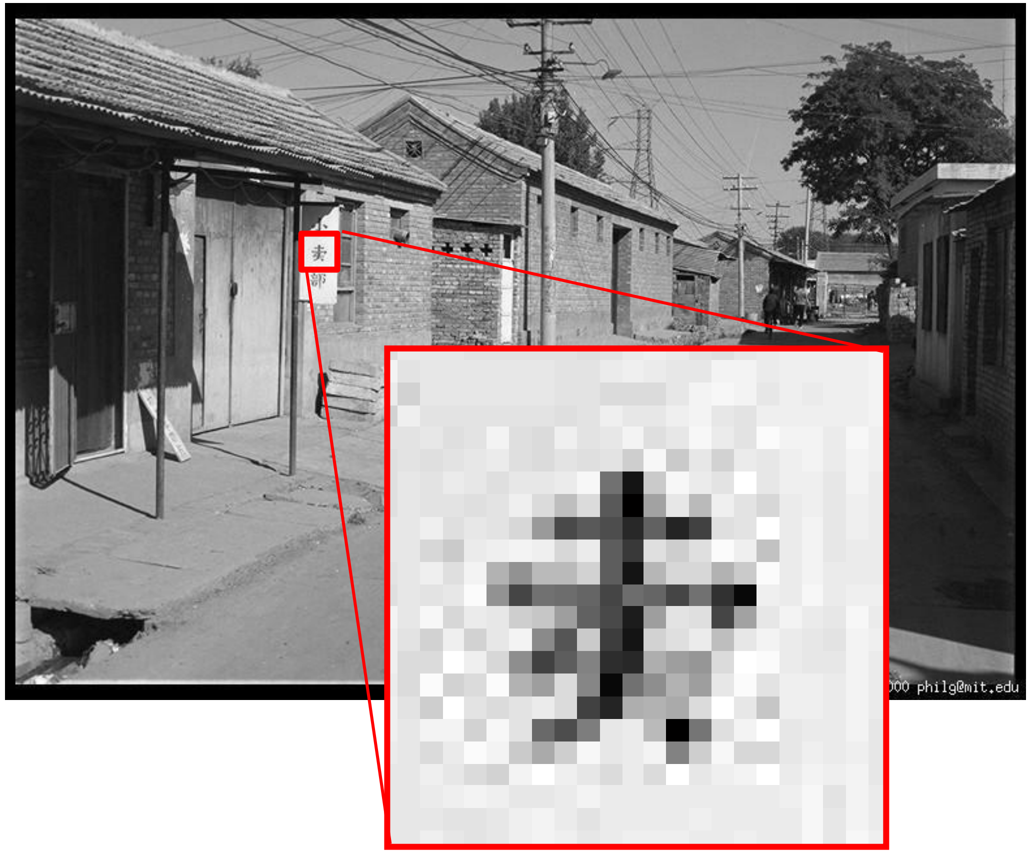
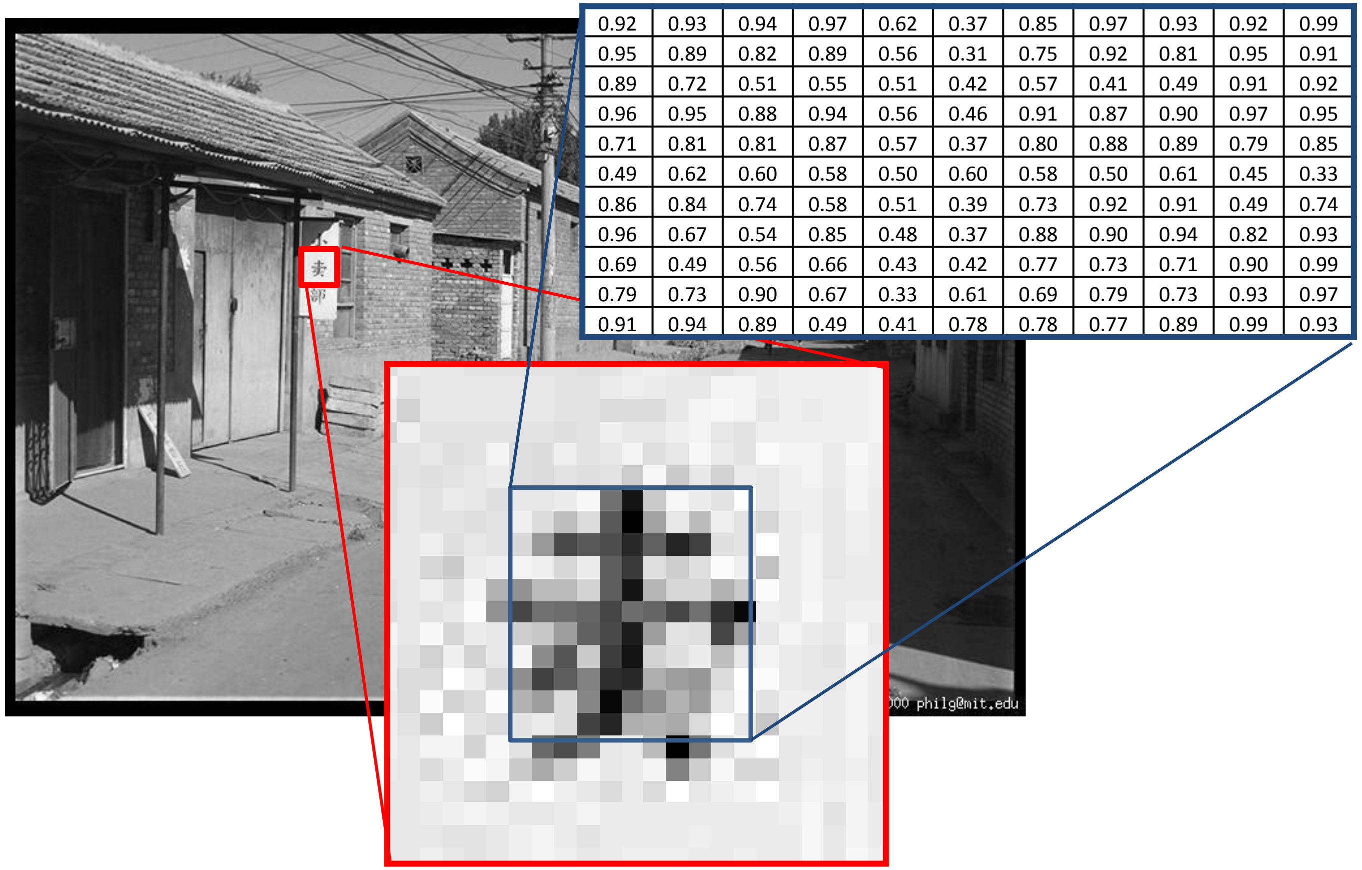


Figure: M. J. Black

Image Representation

- **Digital image:** 2D discrete function f
- **Pixel**: Smallest element of an image f(x,y)



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

Figure: M. J. Black

Sample Problems and Techniques

- Edge Detection
- Image Denoising
- Image Smoothing
- Image Deblurring
- Image Segmentation
- Visual Saliency
- Semantic Segmentation

- PDEs and Variational models
- MRFs
- Graph Theory
- Sparse Coding
- Deep Learning

Image Filtering

- Filtering out the irrelevant information

$$f(x) = u(x) + n(x)$$

$f(x)$	$u(x)$	$n(x)$
↓	↓	↓
observed image	desired image	irrelevant data

- Image denoising, image sharpening, image smoothing, image deblurring, etc.
- Edge detection

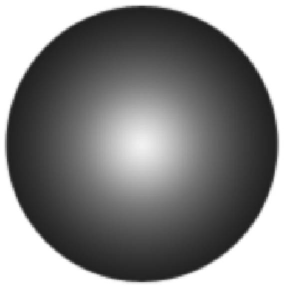
Edge Detection

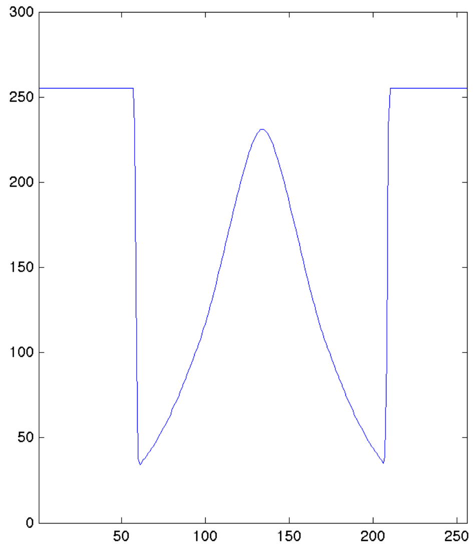
Canny edge detector

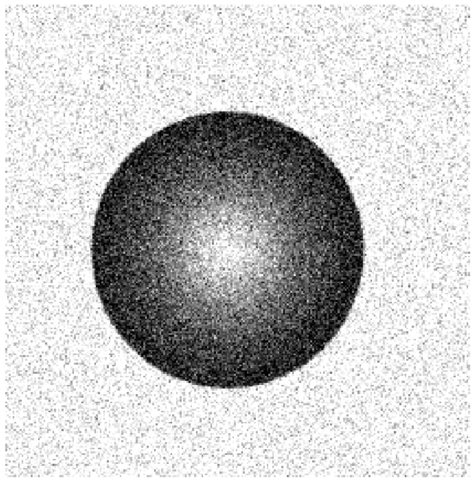
- Edges: abrupt changes in the intensity
  - Uniformity of intensity or color
- Edges to object boundaries

Image Filtering

- <u>Difficulty:</u> Some of the irrelevant image information have characteristics similar to those of important image features







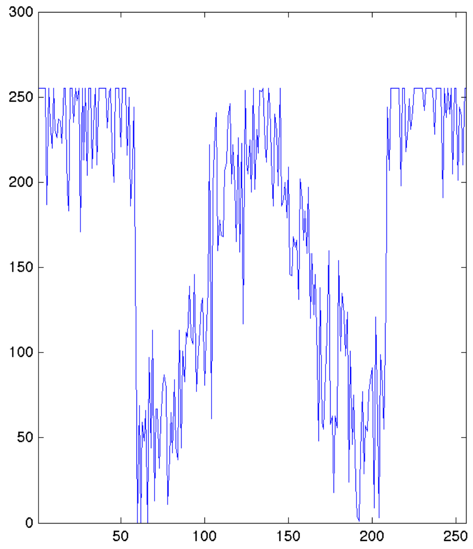


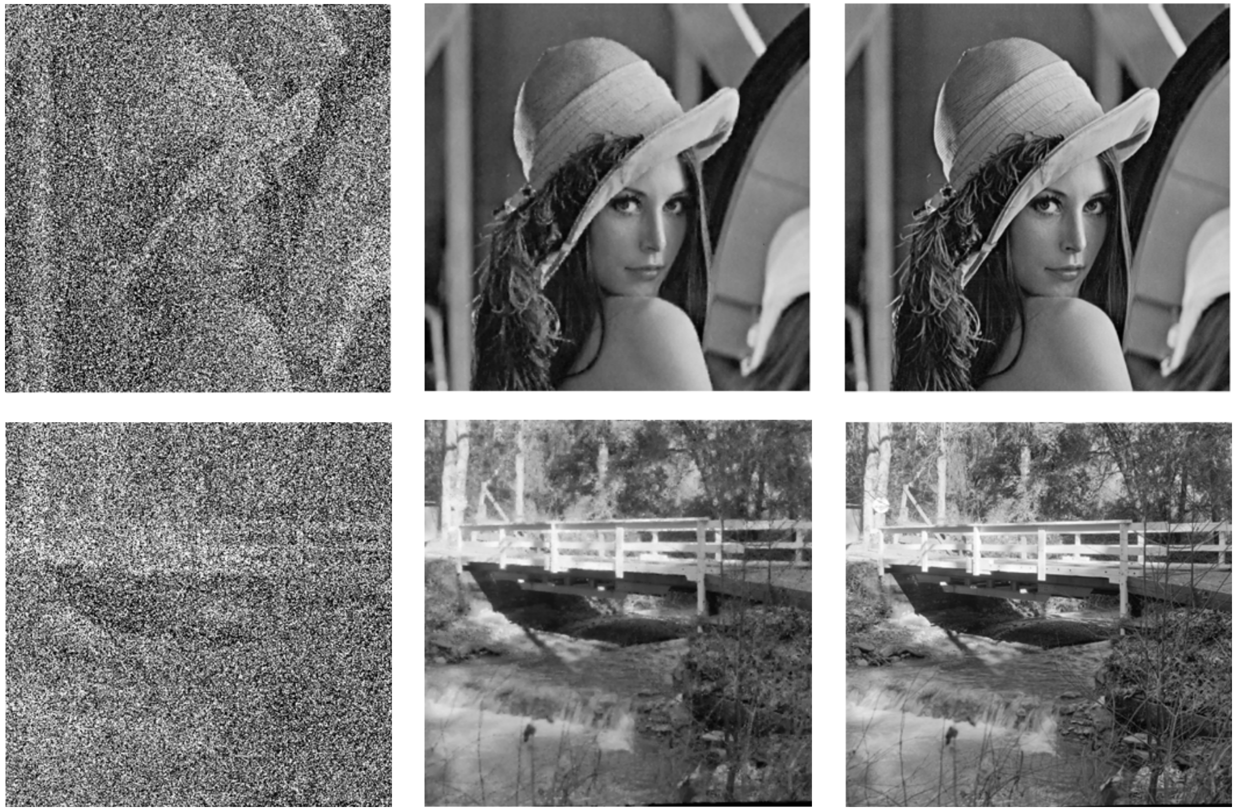
Image Smoothing - A Little Bit of History

- Gaussian Filtering / linear diffusion $\frac{\partial u}{\partial t} = \nabla \cdot (\nabla u) = \nabla^2 u$
  - the most widely used method

- mid 80's – unified formulations
  - methods that combine smoothing and edge detection
  - Geman & Geman'84, Blake & Zisserman'87, Mumford & Shah'89, Perona & Malik'90

Image Denoising

- Images are corrupted with 70% salt-and-pepper noise



Noisy input Recovered image Original image

What do these examples demonstrate?

R. H. Chan, C.-W. Ho, and M. Nikolova, Salt-and-Pepper Noise Removal by Median-Type Noise Detectors and Detail-Preserving Regularization. IEEE TIP 2005

Non-local Means Denoising

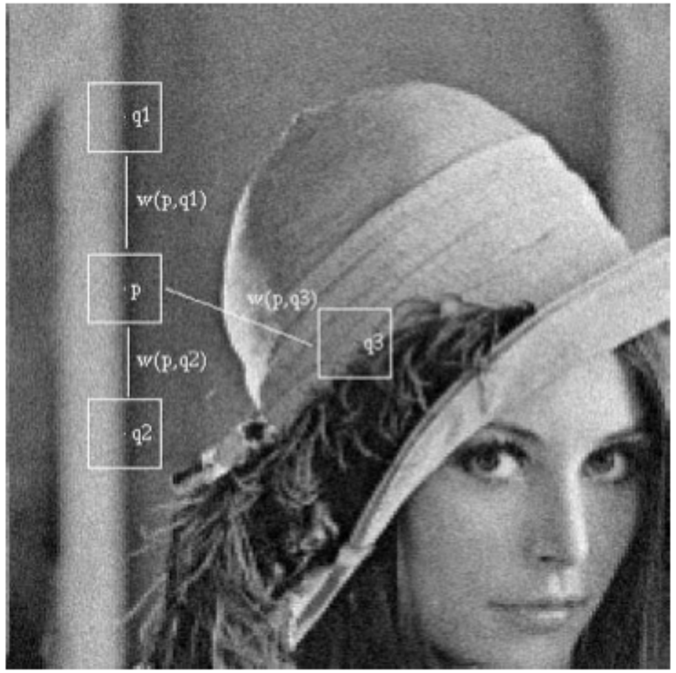
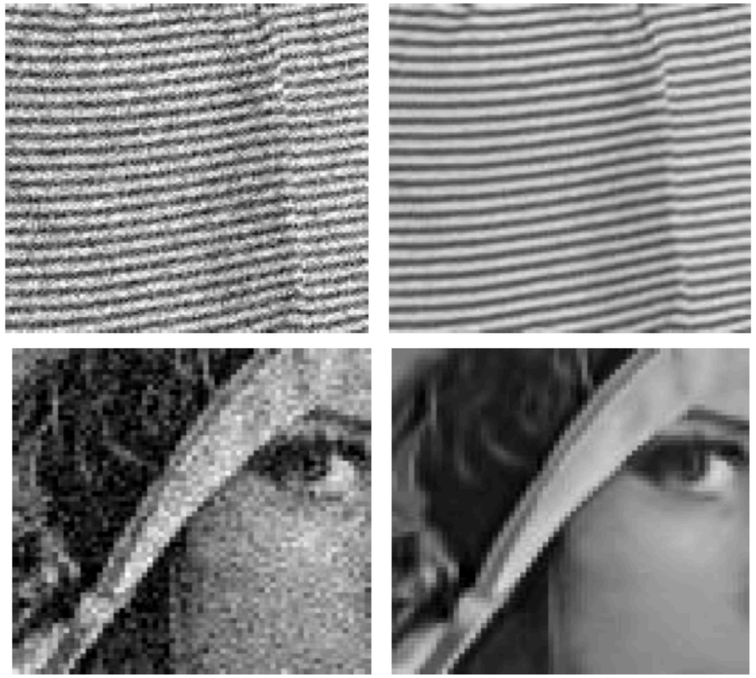


Figure 1. Scheme of NL-means strategy. Similar pixel neighborhoods give a large weight, w(p,q1) and w(p,q2), while much different neighborhoods give a small weight w(p,q3).

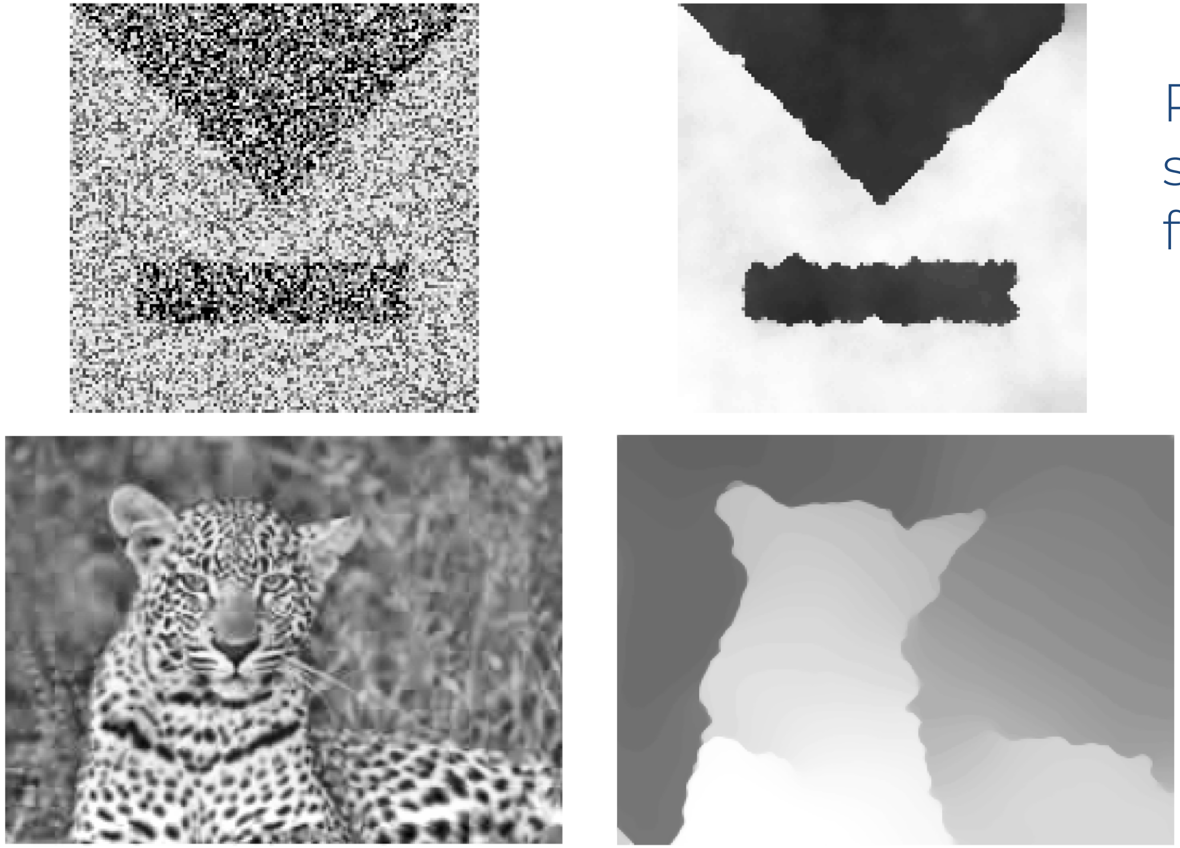


Preserve fine image details and texture during denoising

A. Buades, B. Coll, J. M. Morel, A non-local algorithm for image denoising, CVPR 2005

Context-Guided Filtering

- Use local image context to steer filtering



Preserve main image structures during filtering

E. Erdem and S. Tari, Mumford-Shah Regularizer with Contextual Feedback, JMIV, 2009

Image Smoothing

L. Xu, C. Lu, Y. Xu, J. Jia, Image Smoothing via L0 Gradient Minimization, ACM Trans. Graphics 2011 (SIGGRAPH Asia 2011)

Image Smoothing

L. Karacan, E. Erdem, A. Erdem, Structure Preserving Image Smoothing via Region Covariances, ACM Trans. Graphics 2013 (SIGGRAPH Asia 2013)

Image Deblurring

- Remove blur and restore a sharp image

from a given blurred image

find its latent sharp image

Slide credit: Lee and Cho

Image Deblurring

- Remove blur and restore a sharp image

Input blurred image

Levin et al. CVPR 2010

Slide credit: Lee and Cho

Visual Saliency

- The problem of predicting where people look at images

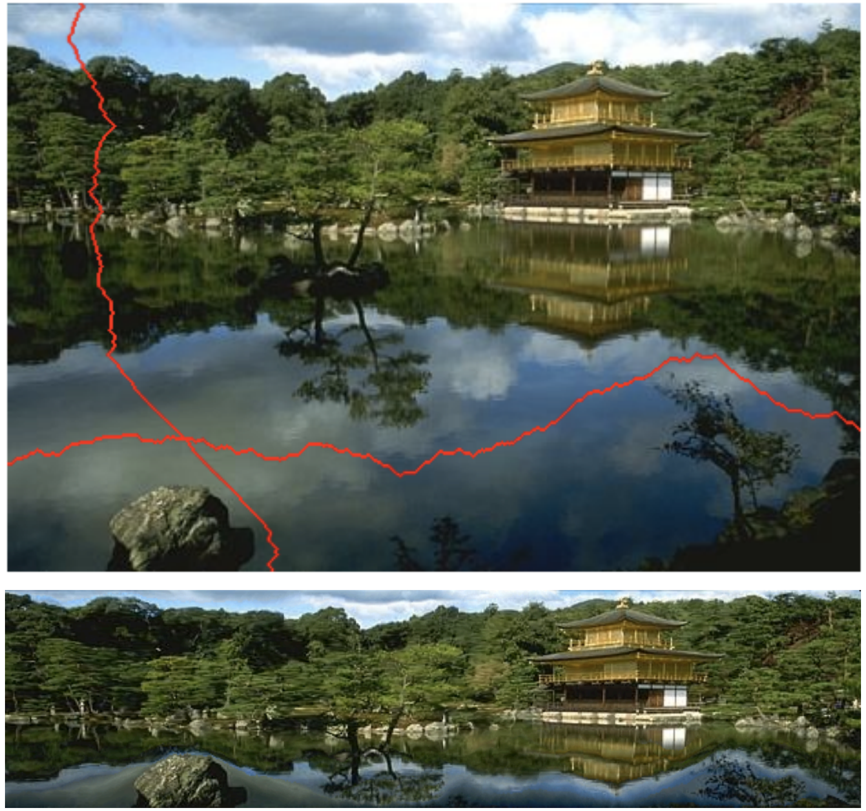
The squares shows where the observers looked in eye tracking experiments

E. Erdem and A. Erdem, Visual saliency estimation by nonlinearly integrating features using region covariances, Journal of Vision 2013

Image Retargetting

- automatically resize an image to arbitrary aspect ratios while preserving important image features

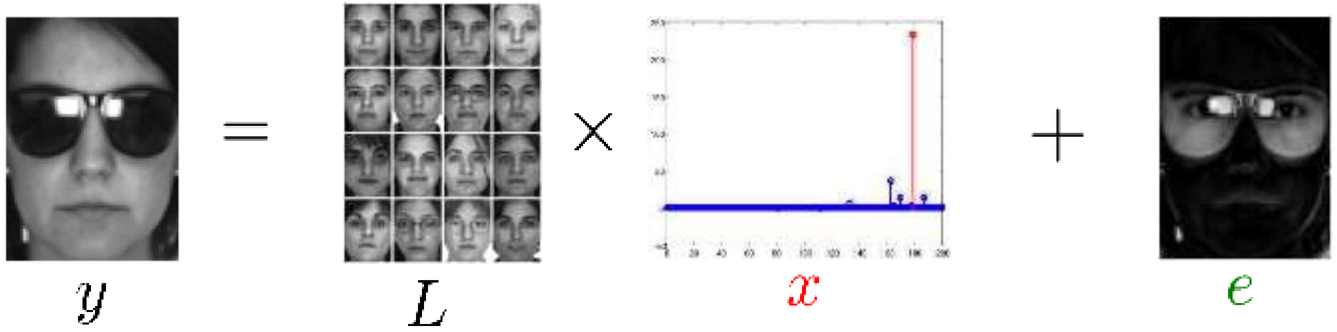
How we define the importance?



S. Avidan and A. Shamir, Seam Carving for Content-Aware Image Resizing, SIGGRAPH, 2007

Sparse Coding

- The problem of finding a small number of representative atoms from a dictionary which when combined with right weights represent a given signal.



$$y = Lx + e$$

Credit: Yi Ma

Image Inpainting

- Reconstructing lost or deteriorated parts of images

What do these examples demonstrate?

M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, Image Inpainting, SIGGRAPH 2000

Image Segmentation

- Partition an image into meaningful regions that are likely to correspond to objects exist in the image

Figures: A. Erdem

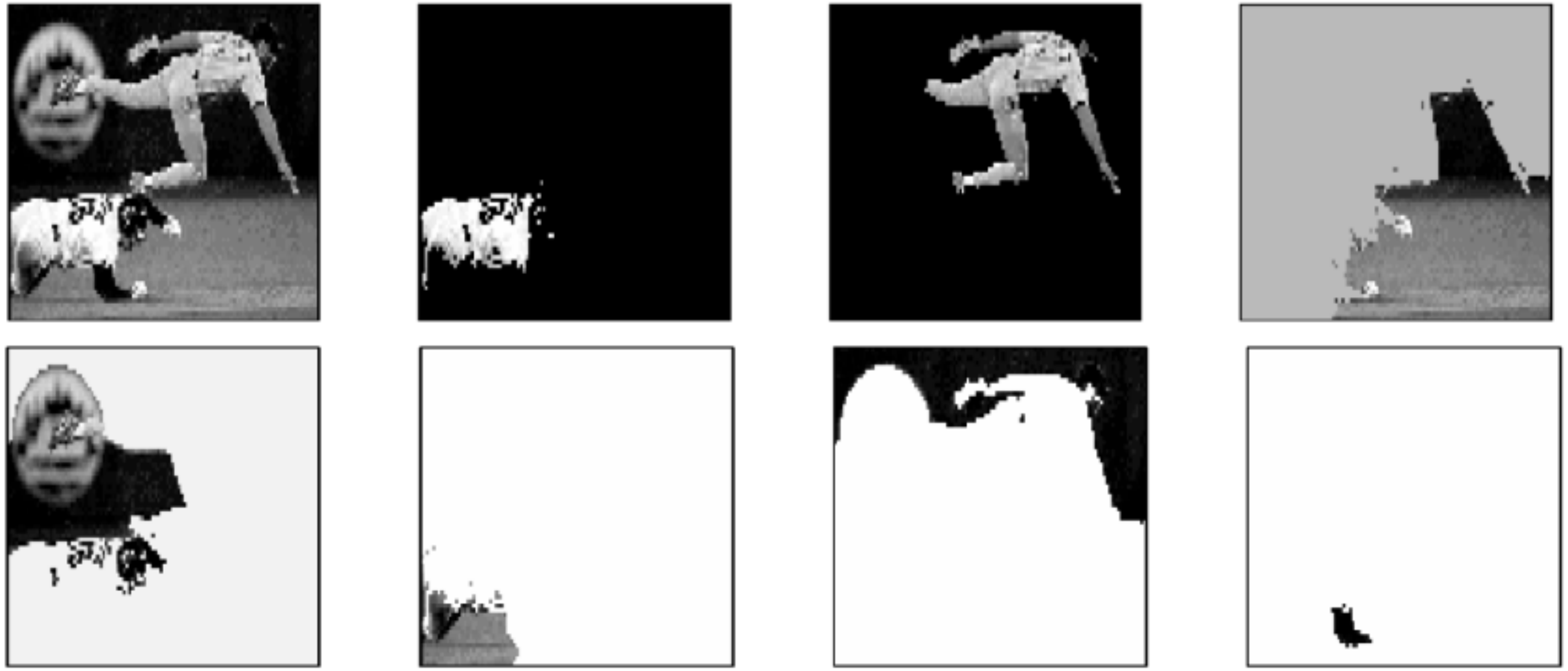
Grouping of pixels

according to what criteria?

high-level object specific knowledge matters!

Normalized Cuts

- A graph-theoretic formulation for segmentation



J. Shi and J. Malik, Normalized Cuts and Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intel., 2001

Normalized Cuts

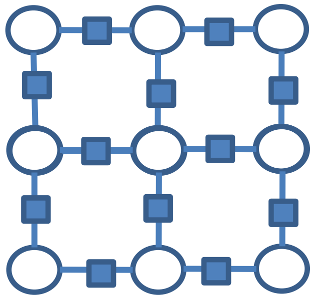
Slide credit: S. Lazebnik

From contours to regions

- gPb-owt-ucm segmentation algorithm

P. Arbelaez, M. Maire, C. Fowlkes and J. Malik, Contour Detection and Hierarchical Image Segmentation,
IEEE Trans Pattern Anal. Mach. Intell. 33(5):898-916, 2011

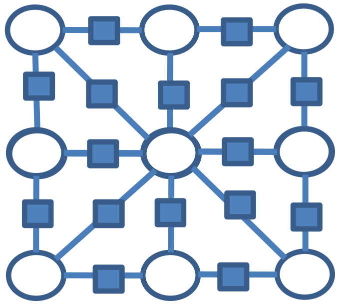
Graphical Models in Vision



4-connected; pairwise MRF

$$E(x) = \sum_{i,j \in N_4} \theta_{ij}(x_i,x_j)$$

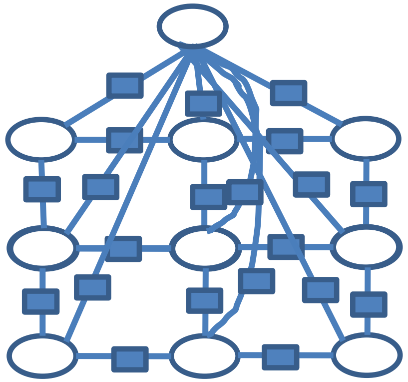
Order 2



higher(8)-connected; pairwise MRF

$$E(x) = \sum_{i,j \in N_8} \theta_{ij}(x_i,x_j)$$

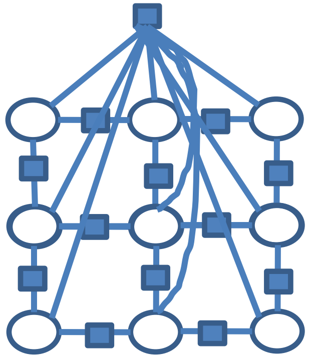
Order 2



MRF with global variables

$$E(x) = \sum_{i,j \in N_8} \theta_{ij}(x_i,x_j)$$

Order 2



Higher-order MRF

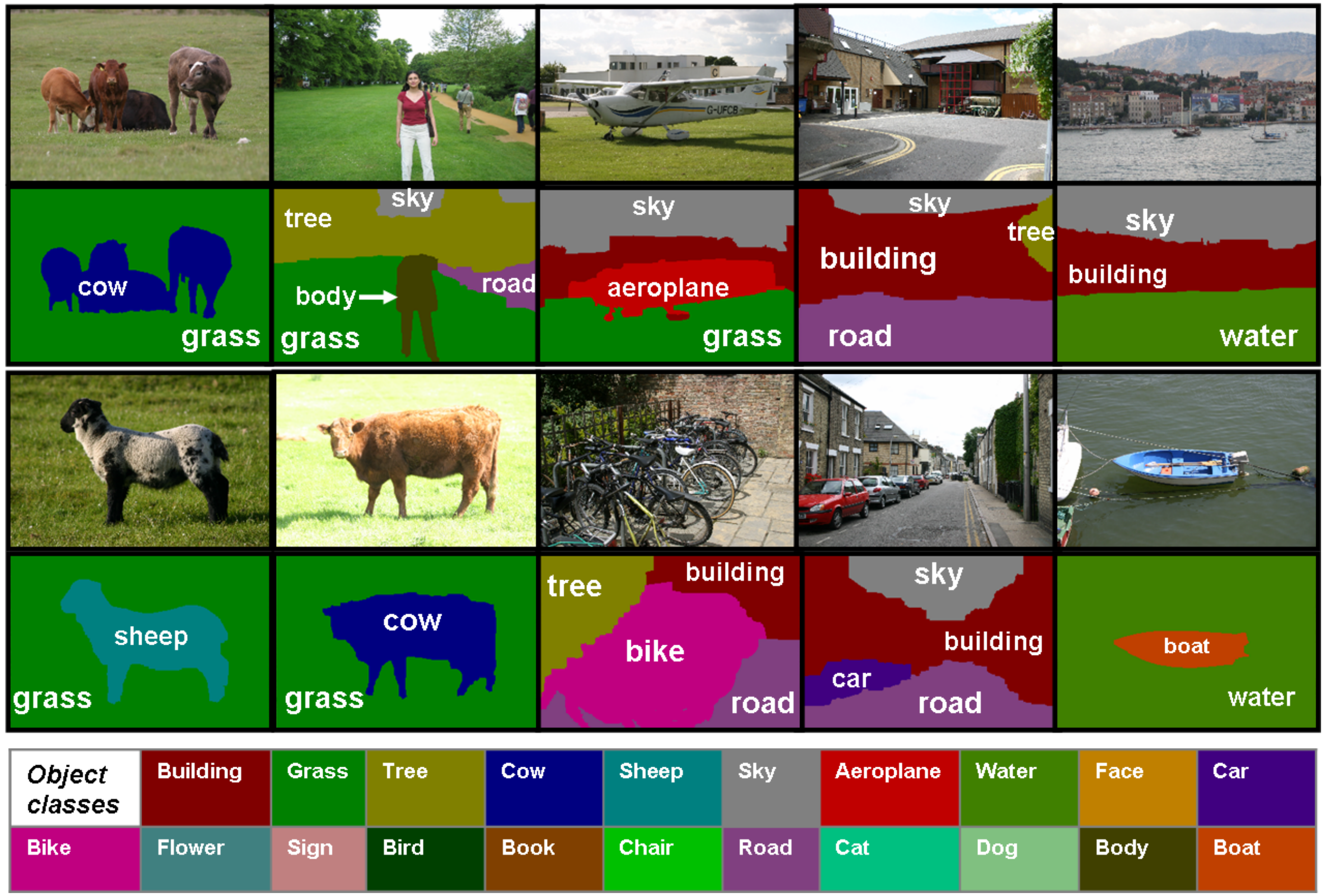
$$E(x) = \sum_{i,j \in N_4} \theta_{ij}(x_i,x_j) + \theta(x_1,\dots,x_n)$$

Order n

C. Rother

Semantic Segmentation

- The problem of joint recognition and segmentation



[TextonBoost; Shotton et al, '06]

C. Rother

Semantic Segmentation

- The problem of joint recognition and segmentation

Carreira et al., Semantic Segmentation with Second-Order Pooling , ECCV 2012

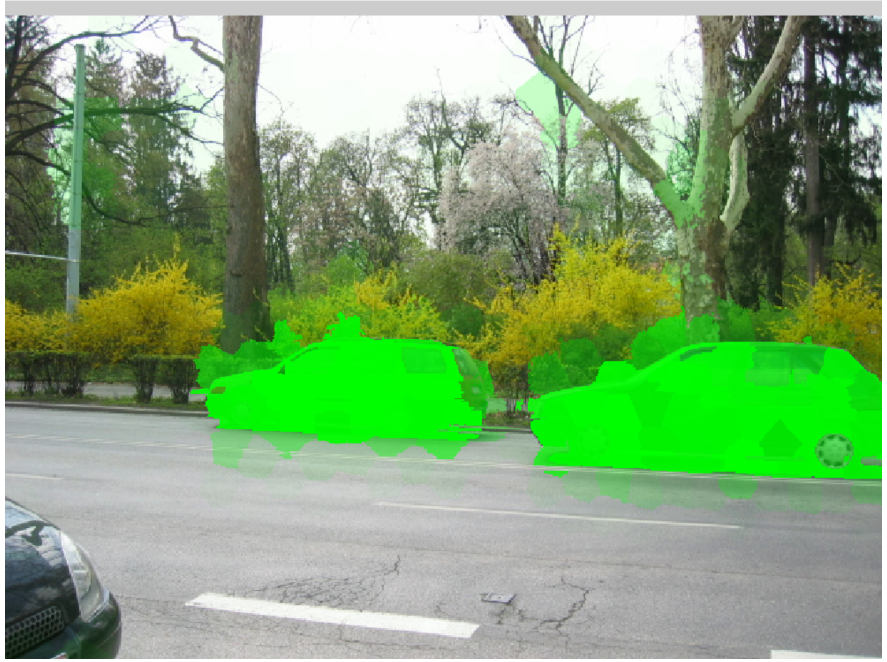
Top-down Saliency

- Task-oriented models (e.g. searching for a target object from a specific category)

A. Kocak et al., Top down saliency estimation via superpixel-based discriminative dictionaries, BMVC 2014

Top-down Saliency

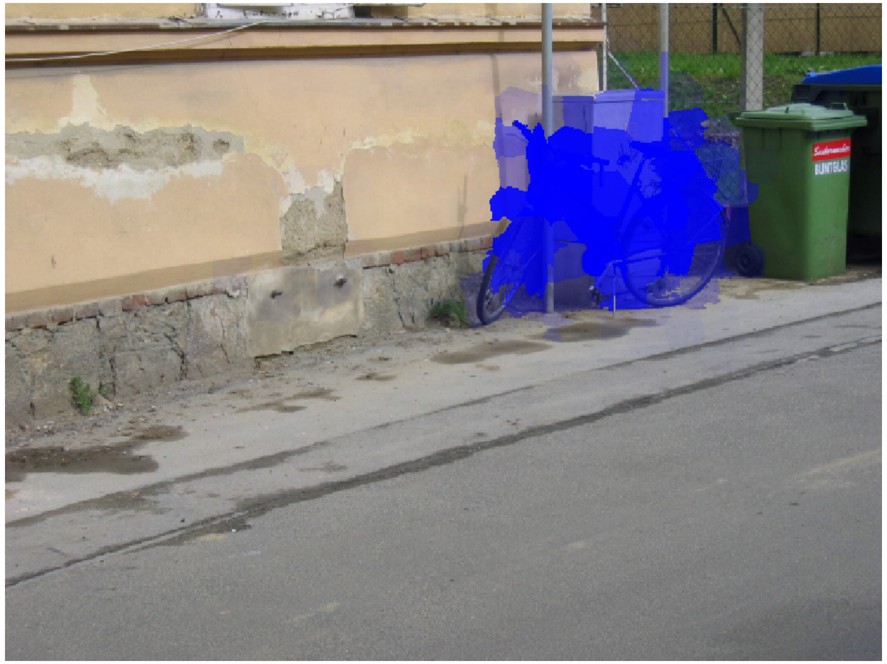
- Task-oriented models (e.g. searching for a target object from a specific category)



A. Kocak et al., Top down saliency estimation via superpixel-based discriminative dictionaries, BMVC 2014

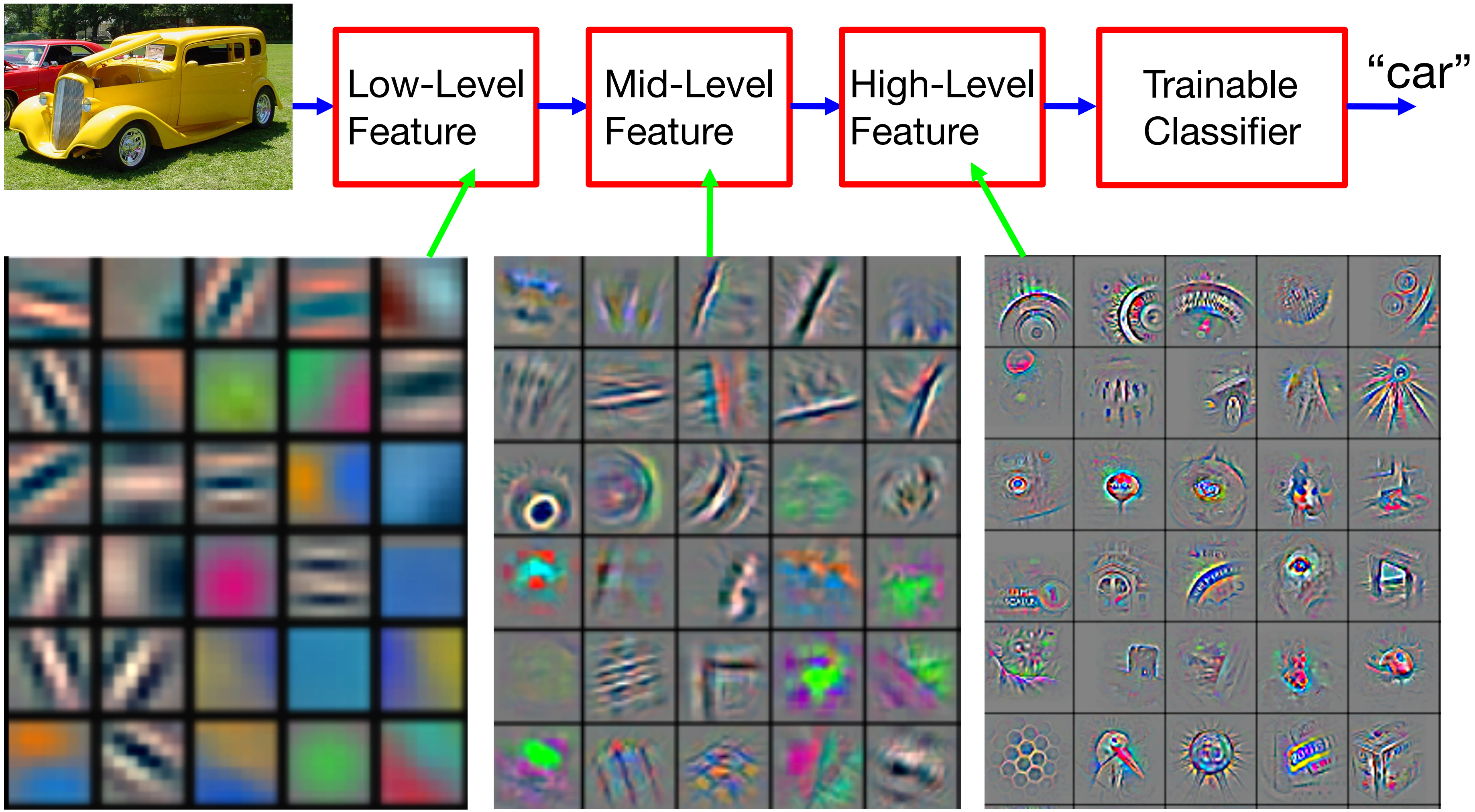
Top-down Saliency

- Task-oriented models (e.g. searching for a target object from a specific category)



A. Kocak et al., Top down saliency estimation via superpixel-based discriminative dictionaries, BMVC 2014

Deep Learning

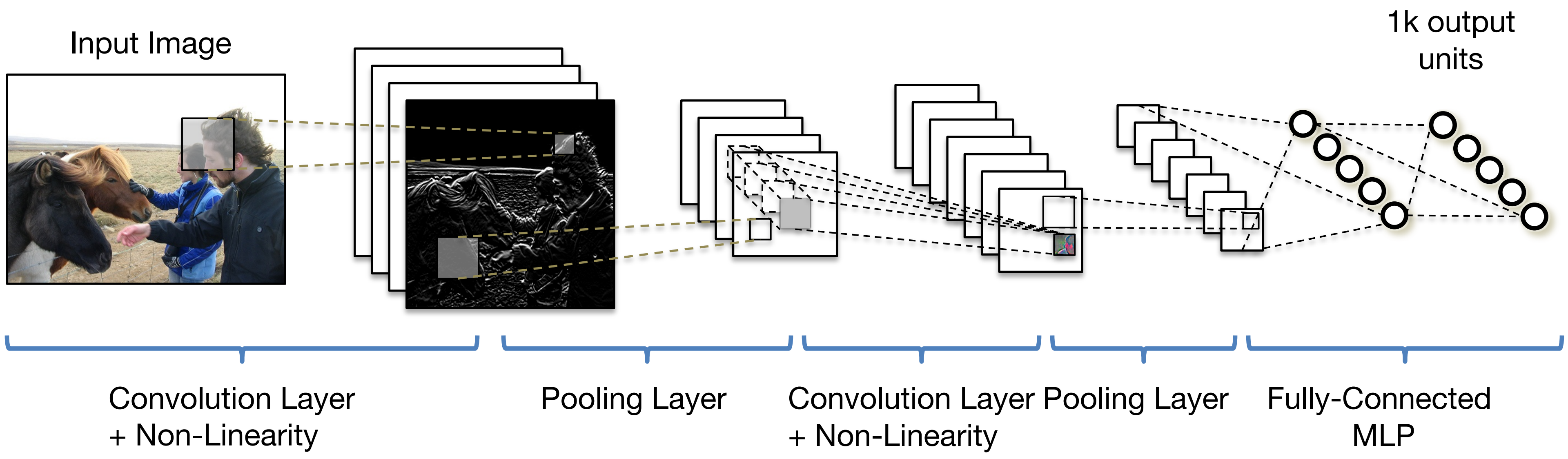


Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

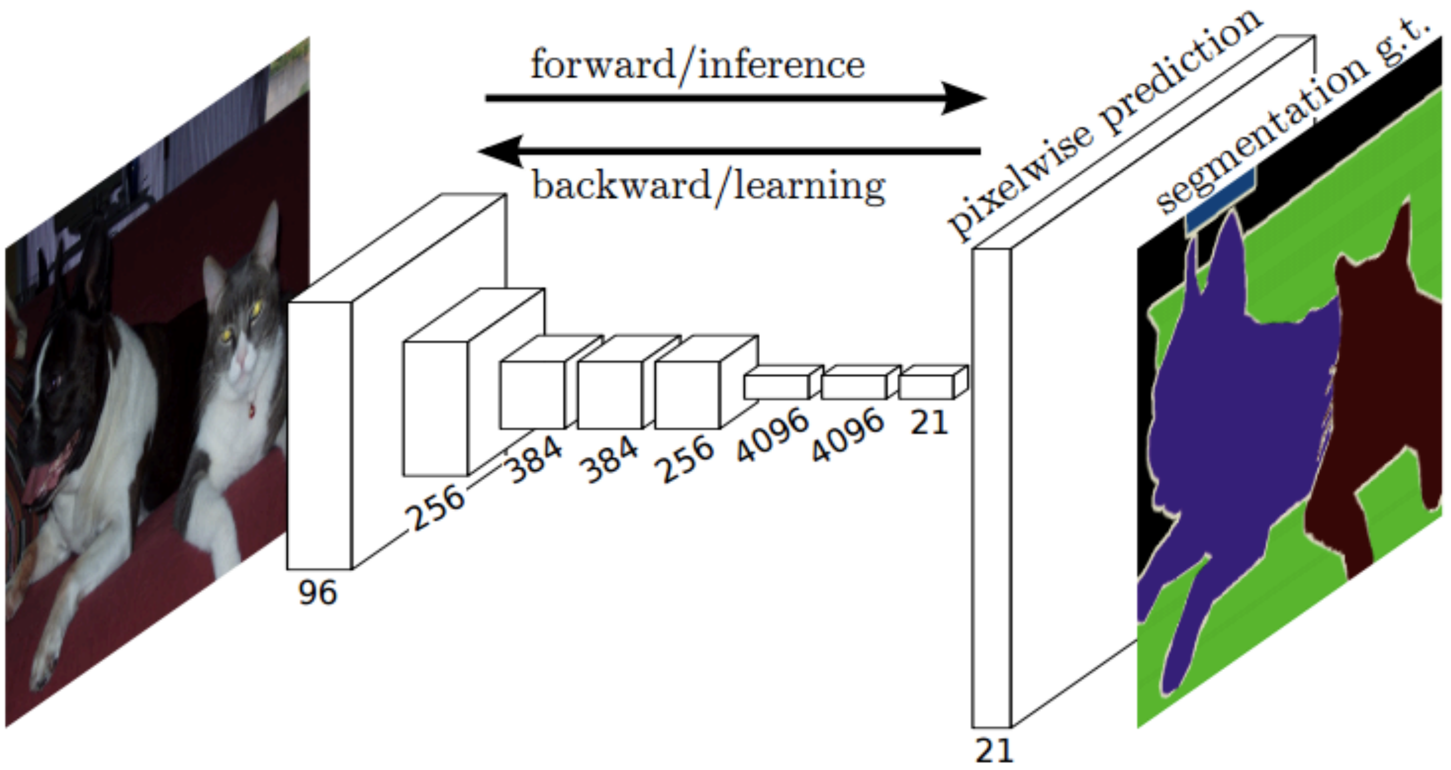
Deep Learning

- [Krizhevsky et al. NIPS12]
  - 54 million parameters; 8 layers (5 conv, 3 fully-connected)
  - Trained on 1.4M images in ImageNet



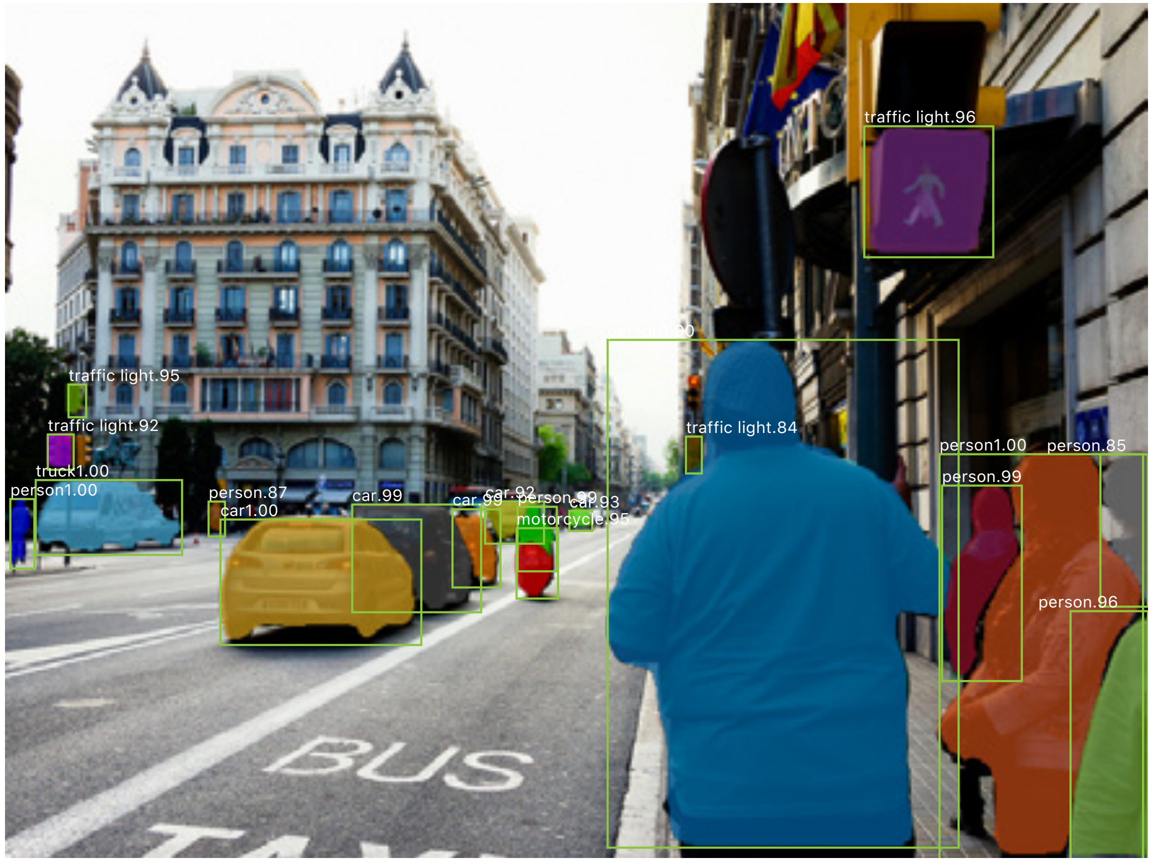
Slide Credit: Dhruv Batra

Semantic Segmentation



Fully Convolutional Networks for Semantic Segmentation [Long, Shelmer & Darrell 2015]

Instance Segmentation



Mask R-CNN [He et al., 2017]

Deep Generative Networks

A Style-Based Generator Architecture for Generative Adversarial Networks [Karras et al., 2018]

Deep Generative Networks

4× SRGAN (proposed)

original

Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [Ledig et al., 2017]

Deep Generative Networks

$4\times$ SRGAN (proposed) original

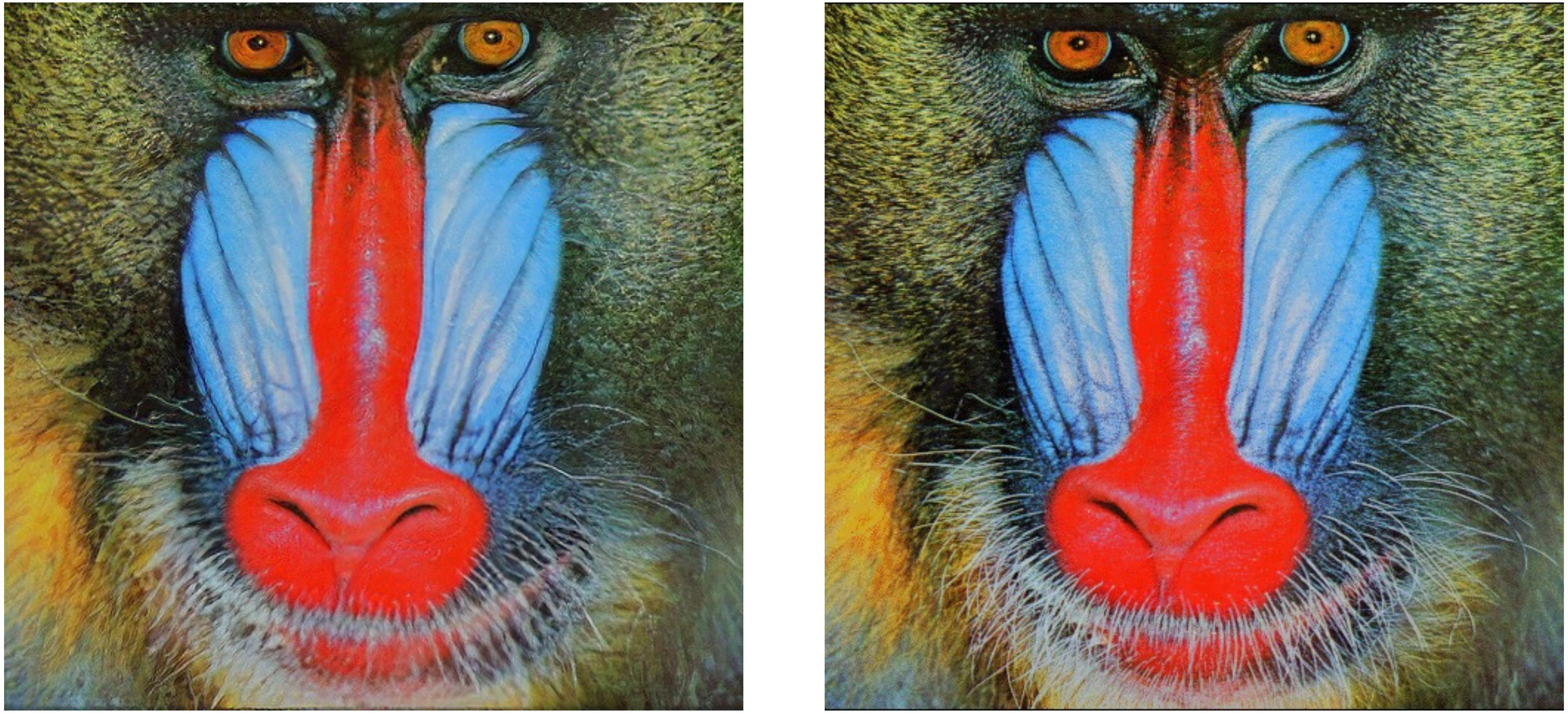


Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [Ledig et al., 2017]

Image to Image Translation

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks [Zhu et al., 2017]

Next lecture

- Linear Filtering,
- Edge/Boundary Detection,
- Image Segmentation