BBM 413 Fundamentals of Image Processing

Erkut Erdem
Dept. of Computer Engineering
Hacettepe University

Introduction

Today

- **Introduction**
  - About the class
  - Organization of this course
- **What is image processing?**
  - What does it mean, to see?
  - Vision as a computational problem
  - Sample image processing problems and applications

Today

- **Introduction**
  - About the class
  - Organization of this course
- **What is image processing?**
  - What does it mean, to see?
  - Vision as a computational problem
  - Sample image processing problems and applications

About this course

- This course is an advanced level undergraduate course about the fundamentals of image processing.
- **Requirements**
  - Programming skills (C/C++, Matlab)
  - Good math background (Calculus, Linear Algebra, Statistical Methods)
  - Little or no prior knowledge of image processing techniques
- **BBM 415 Introduction to Programming Practicum**
  - The students will gain hand-on experience via a set of programming assignments.

About this course (cont'd.)

- **Goals of the course:**
  - to provide an introduction to students who wish to specialize in interrelated disciplines like image processing, computer vision and computational photography
- **Skills to develop:**
  - a foundational understanding and knowledge of concepts that underlie image processing
- **What is image processing?**
  - What does image processing deal with?
  - Computational analysis of low and mid-level vision

BBM 413-415 Team

Instructor

Erkut ERDEM
erkut@cs.hacettepe.edu.tr
Office: 114

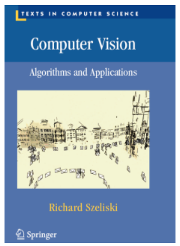
TA

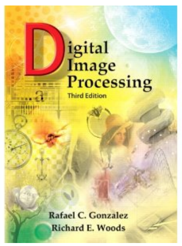
Levent KARACAN
karacan@cs.hacettepe.edu.tr
Office: Vision Lab

- **Office hours:** to be announced!

Textbooks and Reference Material

- Computer Vision: Algorithms and Applications, Richard Szeliski, Springer, 2010
- Digital Image Processing, R. C. Gonzalez, R. E. Woods, 3rd Edition, Prentice Hall, 2008

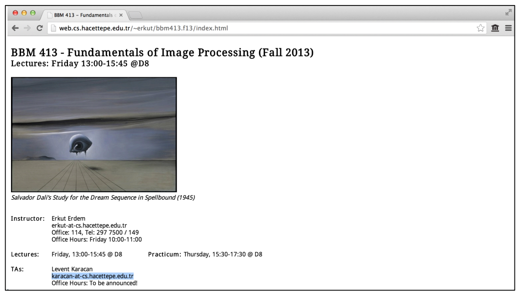




- Lecture notes and handouts
- Papers and journal articles

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/bbm413.f13



Getting Help

- **Office hours**
  - See webpage for the schedule
- **BBM 415 Image Processing Practicum**
  - Course related recitations, practice with example codes, etc.
- **Communication**
  - Announcements and course related discussions through piazza

https://piazza.com/hacettepe.edu.tr/fall2013/bbm413

Course work and grading

- **Reading assignments (15%)**
  - Reading research papers and preparing their summaries
- **Midterm exam (35%)**
  - Closed book and notes
  - In class on November 29th
- **Final exam (45%)**
  - Closed book and notes
  - To be scheduled by Registrar
- **Class participation (5%)**
  - Contribute to Piazza discussions
  - Attend and participate in lectures

Course Overview

- Introduction *(0.5 week)*
- What is image processing? *(0.5 week)*
- Image formation and the digital camera *(1 week)*
- Color perception and color spaces *(1 week)*
- Point operations *(1 week)*
- Spatial filtering *(1 week)*
- Frequency Domain Techniques *(1 week)*
- Image pyramids and wavelets *(1 week)*

Midterm exam

- Gradients, edges, contours *(1 week)*
- Image smoothing *(1 week)*
- Image segmentation *(2 weeks)*
- Advanced topics *(1 week)*

BBM 415 Image Processing Practicum

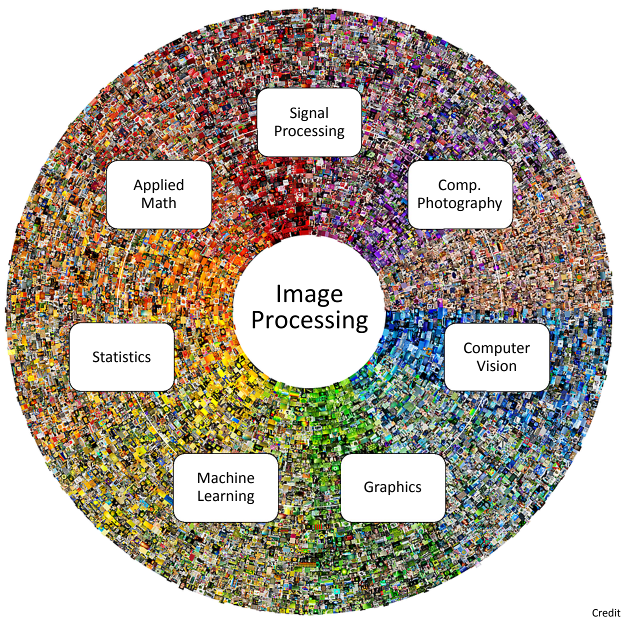
- **Programming assignments (PAs)**
  - Four programming assignments throughout the semester.
  - Each assignment has a well-defined goal such as solving a specific problem.
  - You must work alone on all assignments stated unless otherwise.
- **Important Dates *(Tentative)***
  - PA 1: October 18th
  - PA 2: November 8th
  - PA 3: December 6th
  - PA 4: December 20th

Policies

- **Work groups**
  - You must work alone on all assignments stated unless otherwise
- **Submission**
  - Assignments due at 23:59 on Sunday evenings
  - Electronic submissions (no exceptions!)
- **Lateness penalties**
  - Get penalized **10% per day**
  - No late submission later than **3 days after due date**

Today

- **Introduction**
  - About the class
  - Organization of this course
- **What is image processing?**
  - What does it mean, to see?
  - Vision as a computational problem
  - Sample image processing problems

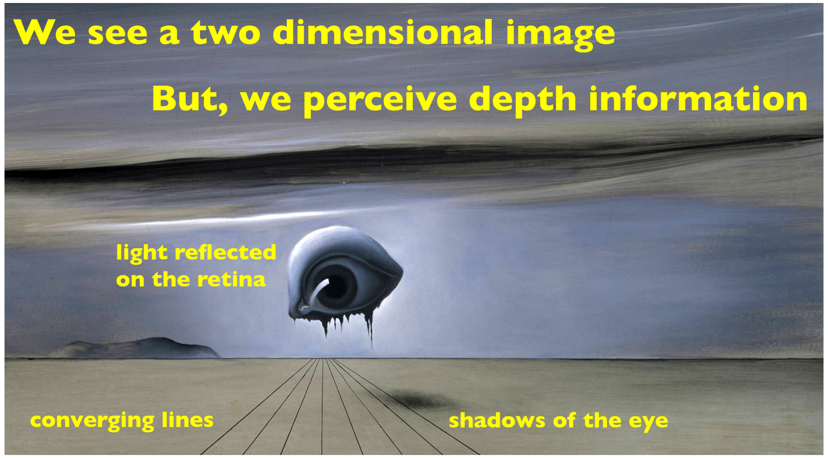


Credit: P. Milanfar

What does it mean, to see?

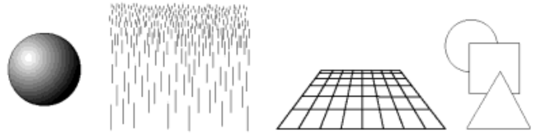
- *"The plain man's answer (and Aristotle's, too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is."* David Marr, Vision, 1982
- Our brain is able to use an image as an input, and interpret it in terms of objects and scene structures.

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



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 ill-posed problems into solvable ones by adding premises: assumptions 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
- The goal of Computer Vision:
  To develop artificial machine vision systems that make inferences related to the scene being viewed through the images acquired with digital cameras.

Things that are easy for us are difficult for computers and viceversa ~ *Marvin Minsky*

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

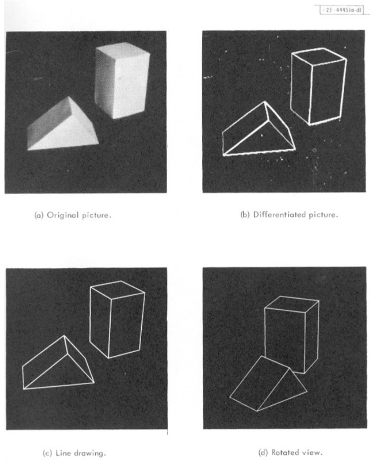
Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Origins of computer vision



L. G. Roberts, *Machine Perception of Three Dimensional Solids,* Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

Slide credit: S. Lazebnik

Vision: a very difficult computational problem, at several levels of understanding

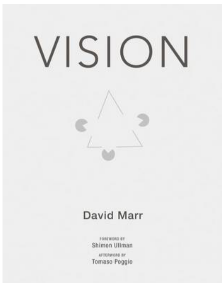
- Vision as an information processing task [David Marr, 1982]

- Three levels of understanding:

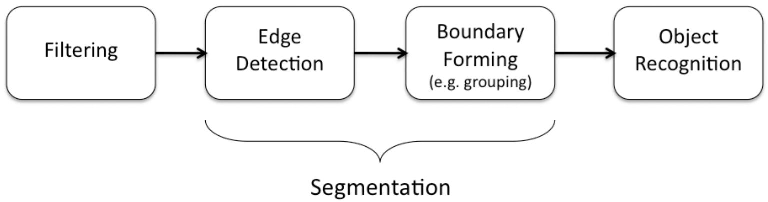
1. Computational theory
   – What is computed? Why it is computed?
2. Representation and Algorithm
   – How it is computed?
   – Input, Output, Transformation
3. Physical Realization
   – Hardware

Reading Assignment #1

- D. Marr (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Chapter 1.
- Due on 25th of October.
- Submit a brief 1-2 pages summary (in English) electronically.
- Use LaTeX to prepare your reports in pdf file format.

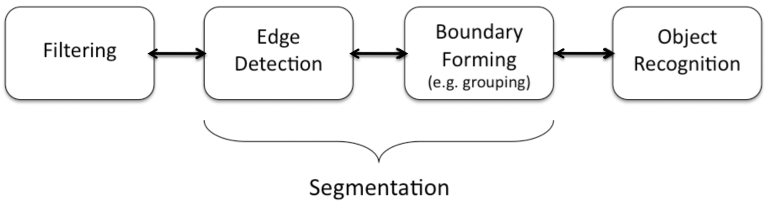


Visual Modules and the Information Flow



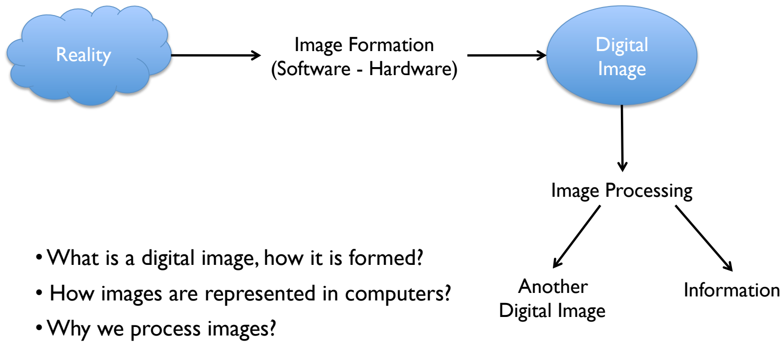
- Visual perception as a data-driven, bottom-up process (traditional view since D. Marr)
- Unidirectional information flow
- Simple low-level cues >> Complex abstract perceptual units

Visual Modules and the Information Flow



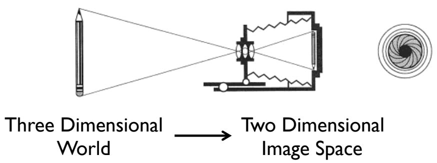
- Vision modules can be categorized into three groups according to their functionality:
  - Low-level vision: filtering out irrelevant image data
  - Mid-level vision: grouping pixels or boundary fragments together
  - High-level vision: complex cognitive processes

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

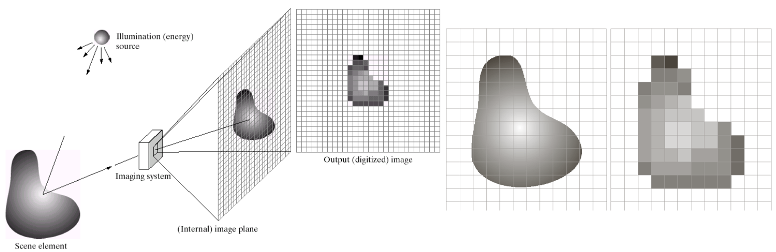


- What is measured in an image location?
  - brightness
  - color

<< 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)

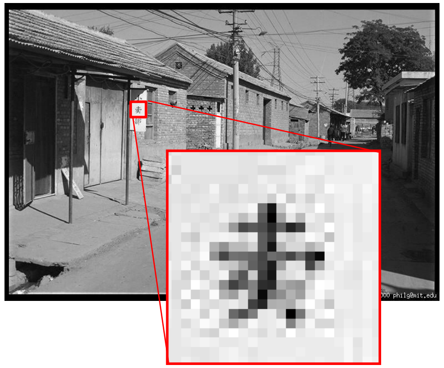
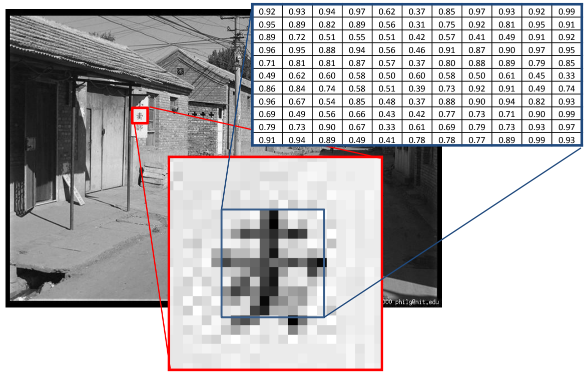


Figure: M. J. Black

Image Representation

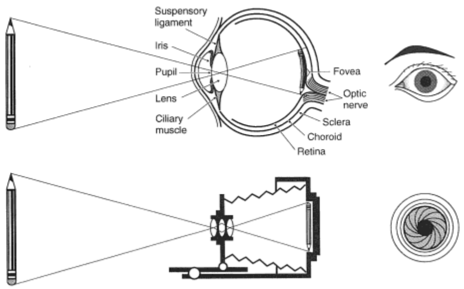
- **Digital image:** 2D discrete function f
- **Pixel**: Smallest element of an image f(x,y)



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

Human Eye

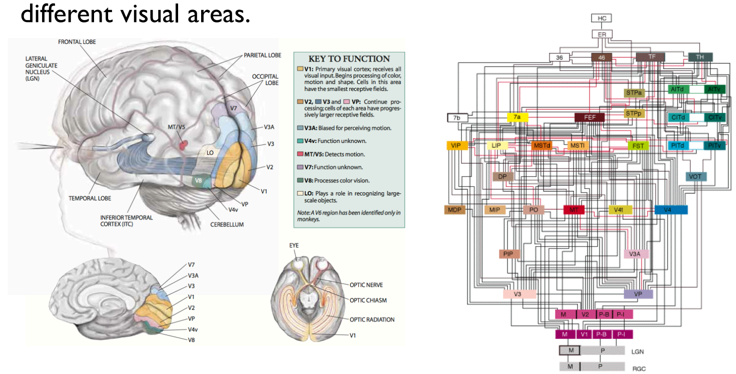


- Two types of receptor cells in retina:
  - Cone Receptor cells: 6-7 million → function in bright light, color sensitive, fine detail
  - Rod receptor cells: 75-150 million → function in dim light, color insensitive, coarse detail
- A recent discovery: Photosensitive retinal ganglion cells → sensitive to blue light

Figures: Gonzalez and Woods, Digital Image Processing, 3rd Edition, 2008
Francis Crick, The Astonishing Hypothesis, 1995

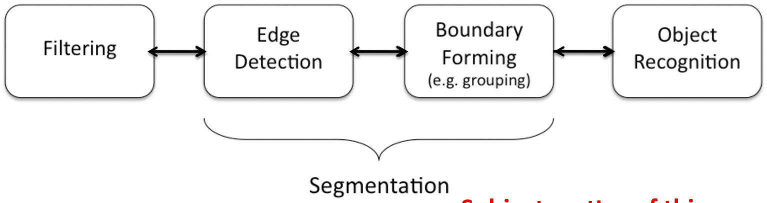
Hierarchy of Visual Areas

- There are many different neural connections between different visual areas.



Figures: Nikos K. Logothetis, Vision: A Window on Consciousness, SciAm, Nov 1999F (on the left)
Felleman & van Essen, 1991 (on the right)

Visual Modules and the Information Flow



Subject matter of this course

- 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

Image Filtering

- Instagram
  - A photo-sharing and social networking service
  - Built-in vintage filters

@ Wikimedia Commons

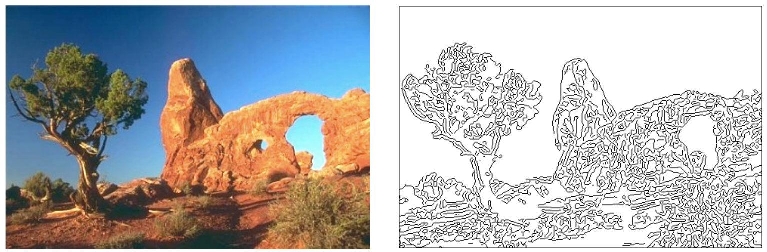
Image Filtering

- Filtering out the irrelevant information

$$\underset{\substack{\downarrow \\ \text{observed} \\ \text{image}}}{f(x)} = \underset{\substack{\downarrow \\ \text{desired} \\ \text{image}}}{u(x)} + \underset{\substack{\downarrow \\ \text{irrelevant} \\ \text{data}}}{n(x)}$$

- Image denoising, image sharpening, image smoothing, image deblurring, etc.
- Edge detection
- Required for many other image image manipulation tasks

Edge Detection



Canny edge detector

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

Image Filtering

- **Difficulty:** Some of the irrelevant image information have characteristics similar to those of important image features

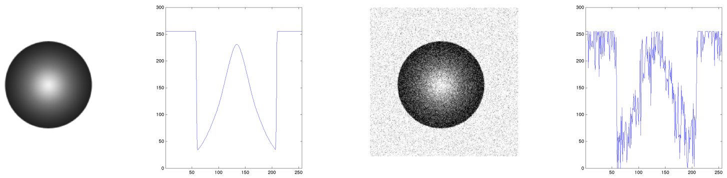


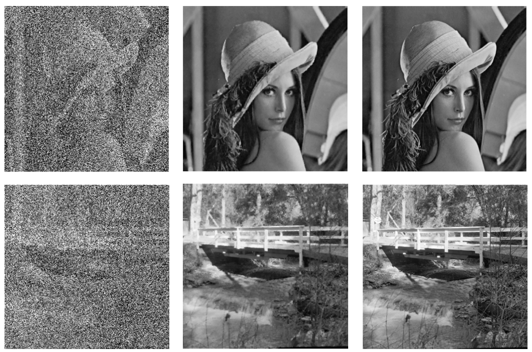
Image Smoothing - A Little Bit of History

- 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

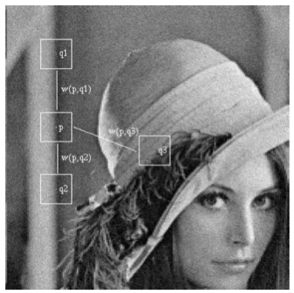
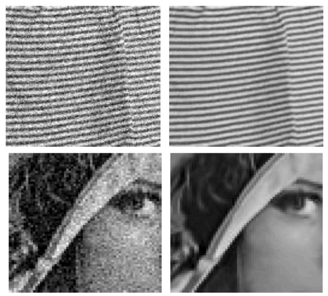


Figure 1. Scheme of NL-means strategy. Similar pixel neighborhoods give a large weight, w(p,q1) and w(p,q2), while much different neighborhoods give a small weight w(p,q3).

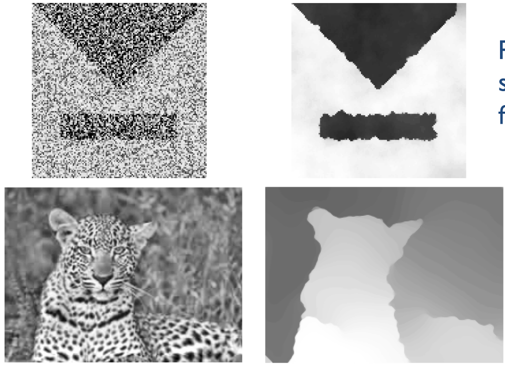


Preserve fine image details and texture during denoising

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

Context-Guided Smoothing

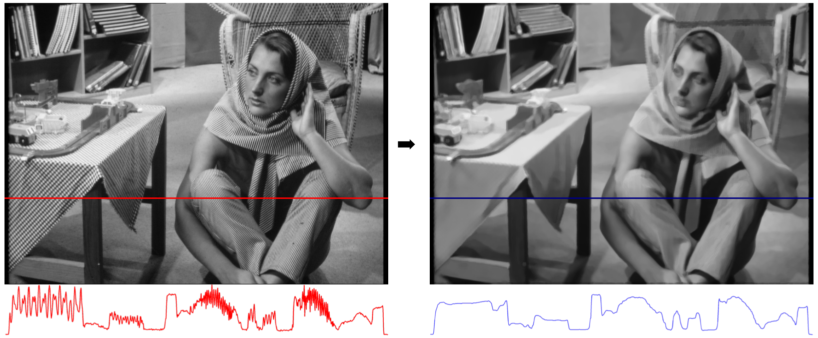
- Use local image context to steer filtering



Preserve main image structures during filtering

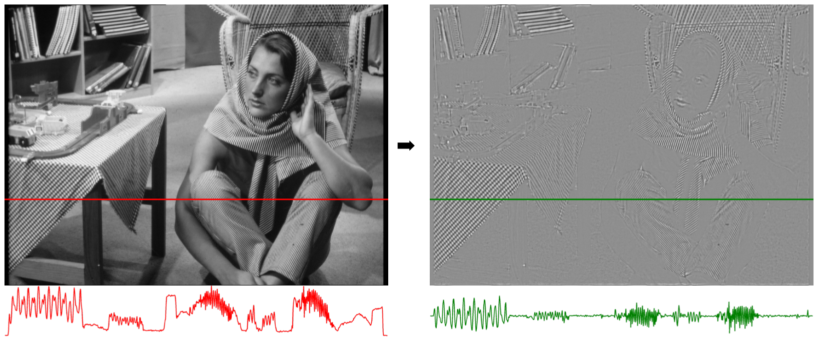
E. Erdem and S. Tari, Mumford-Shah Regularizer with Contextual Feedback, JMIV, 2009

Structure-Preserving Smoothing



L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2012

Structure-Preserving Smoothing



L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2012

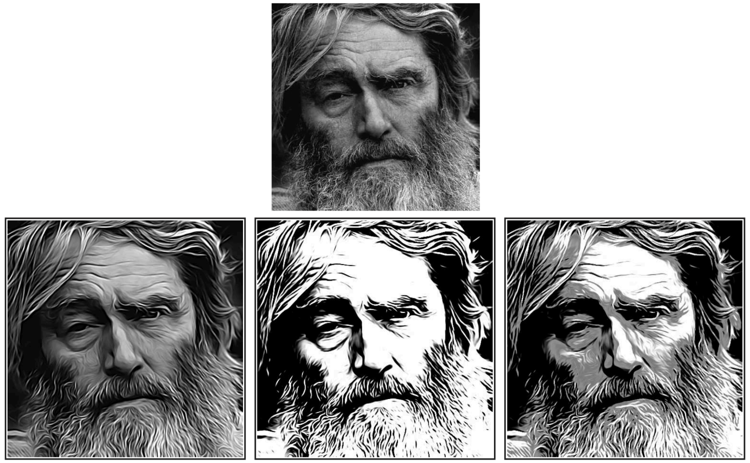
Image Abstraction

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2012

Detail Enhancement

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2012

Artistic Stylizations



H. Winnemöller, J. E. Kyprianidis and S. C. Olsen, XDoG: An eXtended difference-of-Gaussians compendium including advanced image stylization, Computers & Graphics, 2012

Image Segmentation

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

Grouping of pixels

according to what criteria?

high-level object specific knowledge matters!

Figures: A. Erdem

Image Segmentation

- Boundary-based segmentation
- Region-based segmentation
- Unified formulations

Snakes

- Curve Evolution - parametric curve formulation

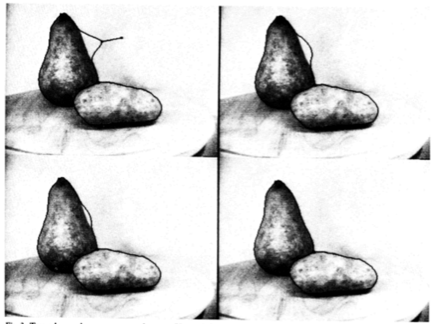
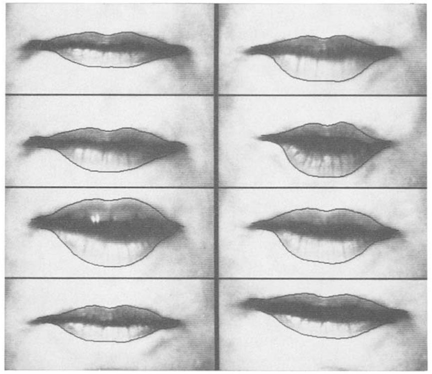


Fig. 3. Two edge snakes on a pear and potato. Upper-left: The user has pulled one of the snakes away from the edge of the pear. Others: After the user lets go, the snake snaps back to the edge of the pear.

M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active Contour Models, IJCV, 1988

Snakes

- Curve Evolution - parametric curve formulation

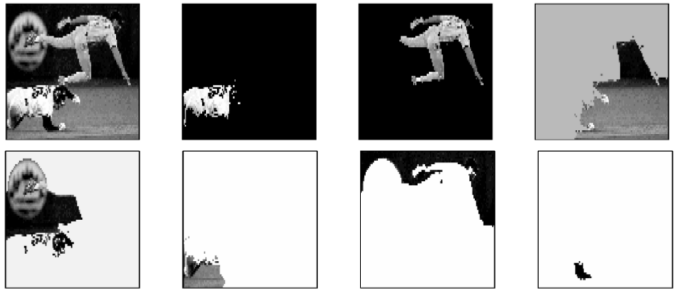


Non-rigid, deformable objects can change their shape over time, e.g. lips, hands…

M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active Contour Models, IJCV, 1988

Normalized Cuts

- A graph-theoretic formulation for segmentation



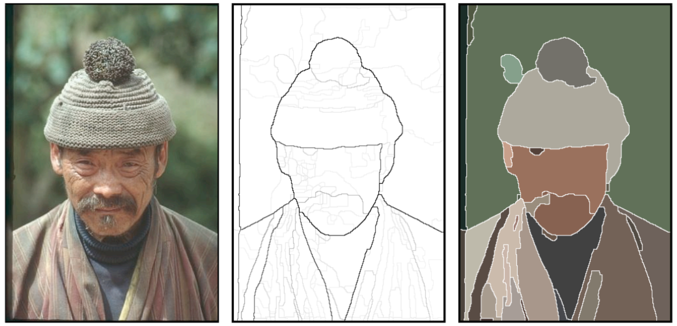
J. Shi and J. Malik, Normalized Cuts and Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intel.

Normalized Cuts

Slide credit: S. Lazebnik

From contours to regions

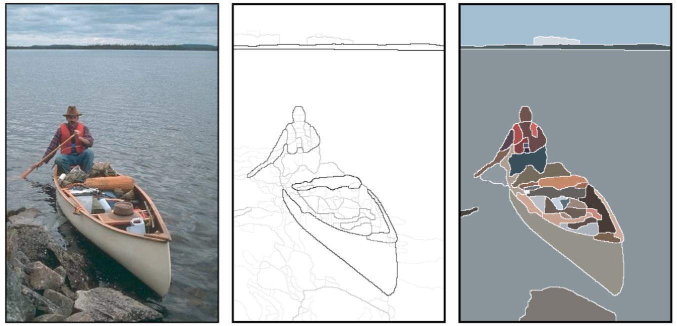
- State-of-the-art: gPb-owt-ucm segmentation algorithm



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

From contours to regions

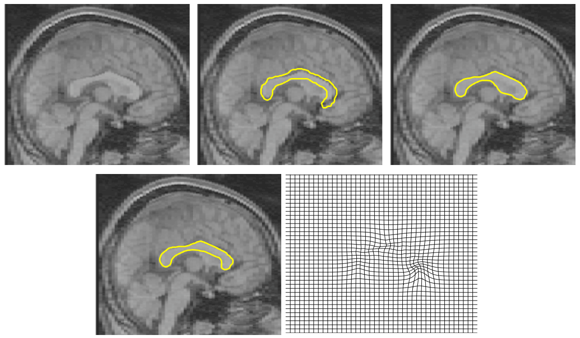
- State-of-the-art: gPb-owt-ucm segmentation algorithm



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

Prior-Shape Guided Segmentation

- Incorporate prior shape information into the segmentation process



Our result Deformation map

E. Erdem, S. Tari, and L. Vese, Segmentation Using The Edge Strength Function as a Shape Prior within a Local Deformation Model, ICIP 2009

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 Resizing

- Resize an image to arbitrary aspect ratios

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

Image Retargeting

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

Image Retargeting

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2012

Next week

- Image formation
- Digital camera and images