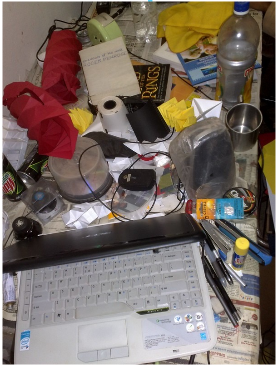
Computer Vision

Lecture 1
Course Introduction

Part 1

What is Computer Vision?



Where is the gluestick? Find the book - what's its full title?

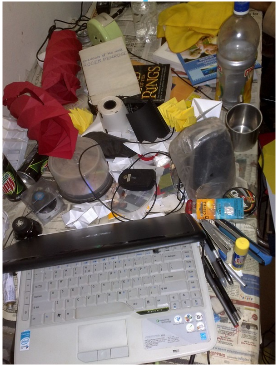
Credit: Bharath Kishore, Flickr CC License

What is wrong with this image?

Credit: Erik Johansson

[Computer Vision: What and Why]

What is Computer Vision?



Where is the gluestick? Find the book - what's its full title?

Credit: Bharath Kishore, Flickr CC License

What is wrong with this image?

Credit: Erik Johansson

Can a machine answer the above questions?

What is Computer Vision?

Computer Vision

A field that seeks to automate and endow a computing framework with the ability to interpret images the way humans do.

A sub-topic of Artificial Intelligence.

What is Computer Vision?

Computer Vision

A field that seeks to automate and endow a computing framework with the ability to interpret images the way humans do.

A sub-topic of Artificial Intelligence.

Other Definitions

"the construction of explicit, meaningful descriptions of physical objects from images" (Ballard & Brown, 1982)

"computing properties of the 3D world from one or more digital images" (Trucco & Verri, 1998)

"to make useful decisions about real physical objects and scenes based on sensed images" (Sockman & Shapiro, 2001)

Why? Applications of Computer Vision

Autonomous Vehicles

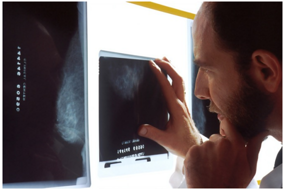
Credit: smoothgrover22, Flickr CC License

Surveillance

Credit: Yeong Nam, Flickr CC License

Factory Automation

Credit: KUKA Roboter GmbH, Bachmann



Medical Imaging

Credit: National Cancer Institute

Human-Computer Interaction

Credit: Vancouver Film School

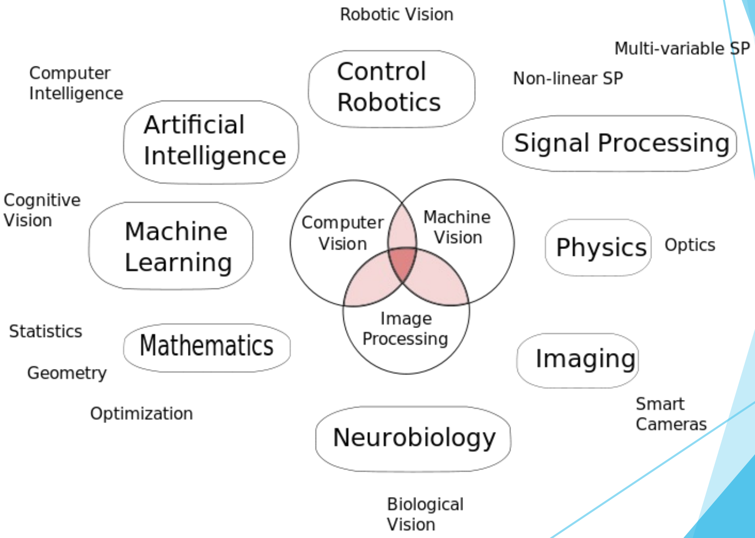
Visual Effects

Credit: AntMan3001, Flickr CC License

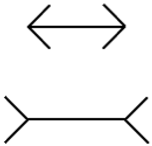
Applications of Computer Vision: More…

- **Retail and Retail Security** (Amazon Go, Virtual Try-on, StopLift)
- **Healthcare** (Blood Loss Detector, DermLens)
- **Agriculture** (SlantRange, Cainthus - Livestock facial recognition)
- **Banking and Finance** (Mobile Deposit, Insurance Risk Profiling)
- **Remote Sensing** (Land Use Understanding, Forestry Modeling)
- **Structural Health Monitoring** (Oilwell Inspection, Drone-based Bridge Inspection and 3D Reconstruction)
- **Document Understanding** (Optical Character Recognition, Robotic Process Automation)
- **Tele- and Social Media** (Image Understanding, Brand Exposure Analytics)
- **Augmented Reality** (TechSee Visual Support, Warehouse and Enterprise Management)

Perspectives of Study



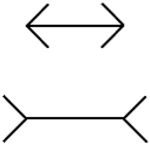
Why is it hard?[1]



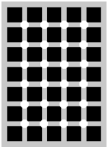
Müller-Lyer illusion: Which line is longer?

[1]Credit: Szeliski, Computer Vision: Algorithms and Applications, 2010

Why is it hard?[1]



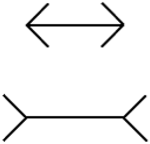
Müller-Lyer illusion: Which line is longer?



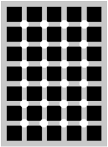
Variation of Hermann grid illusion: What do you see at the intersections?

[1]Credit: Szeliski, Computer Vision: Algorithms and Applications, 2010

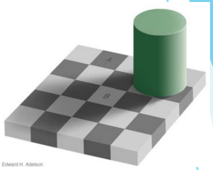
Why is it hard?[1]



Müller-Lyer illusion: Which line is longer?



Variation of Hermann grid illusion: What do you see at the intersections?

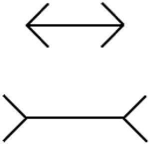


Adelson's brightness constancy illusion: Which is brighter, A or B?

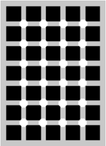
---

[1]Credit: Szeliski, Computer Vision: Algorithms and Applications, 2010

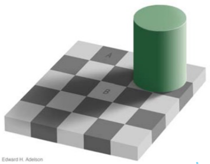
Why is it hard?[1]



Müller-Lyer illusion: Which line is longer?



Variation of Hermann grid illusion: What do you see at the intersections?



Adelson's brightness constancy illusion: Which is brighter, A or B?

Count the red Xs in both figures, which is harder?

[1]Credit: Szeliski, Computer Vision: Algorithms and Applications, 2010

Why is it hard?

- Many practical use cases are **inverse model** applications
  - No knowledge of how an image was taken or camera parameters - but need to model the real world in which picture/video was taken (shape, lighting, color, objects, interactions). $\Longrightarrow$ Need to almost always model from incomplete/partial noisy information
  - *Forward models* are used in physics (radiometry, optics, and sensor design) and in computer graphics

Why is it hard?

- Many practical use cases are **inverse model** applications
  - No knowledge of how an image was taken or camera parameters - but need to model the real world in which picture/video was taken (shape, lighting, color, objects, interactions). $\Longrightarrow$ Need to almost always model from incomplete/partial noisy information
  - *Forward models* are used in physics (radiometry, optics, and sensor design) and in computer graphics
- High-dimensional data $\Longrightarrow$ heavy computational requirements

Why is it hard?

- Many practical use cases are **inverse model** applications
  - No knowledge of how an image was taken or camera parameters - but need to model the real world in which picture/video was taken (shape, lighting, color, objects, interactions). $\Longrightarrow$ Need to almost always model from incomplete/partial noisy information
  - *Forward models* are used in physics (radiometry, optics, and sensor design) and in computer graphics
- High-dimensional data $\Longrightarrow$ heavy computational requirements
- Computer vision is AI-complete

Credit: Anish Chopra, Medium.com

Why is it hard?

- No complete models of the human visual system exist
  - Existing models largely related to subsystems, not holistic
  - What is perceived, and what is cognized? When is an object important for a task, and when is the context important?

Why is it hard?

- No complete models of the human visual system exist
  - Existing models largely related to subsystems, not holistic
  - What is perceived, and what is cognized? When is an object important for a task, and when is the context important?
- Verifiability of mathematical/physical models non-trivial
  - How should similarity/dissimilarity between representations be defined? Is this a distance metric? Do all images follow such a distance metric?
  - How would a manipulation (counterfactual) in a given (potentially noisy) environment behave, w.r.t. the captured image/video? Can a physical model capture this?

Credit: Anish Chopra, Medium.com

Computer Vision: Topics

Learning-based Vision	Visual Recognition, Detection, Segmentation, Tracking, Retrieval, etc
Geometry-based Vision	Feature-based Alignment, Image Stitching, Epipolar Geometry, Structure from Motion, 3D Reconstruction, etc
Physics-based Vision	Computational Photography, Photometry, Light-fields, Color Spaces, Shape-from-X, Reflection, Refraction, Polarization, Diffraction, Interference, etc

Computer Vision: Topics

Learning-based Vision

Visual Recognition, Detection, Segmentation, Tracking, Retrieval, etc

Focus of this course

Geometry-based Vision

Feature-based Alignment, Image Stitching, Epipolar Geometry, Structure from Motion, 3D Reconstruction, etc

Physics-based Vision

Computational Photography, Photometry, Light-fields, Color Spaces, Shape-from-X, Reflection, Refraction, Polarization, Diffraction, Interference, etc

Part 2

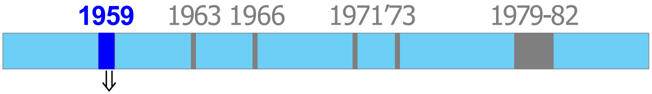
- **History**

Part 2

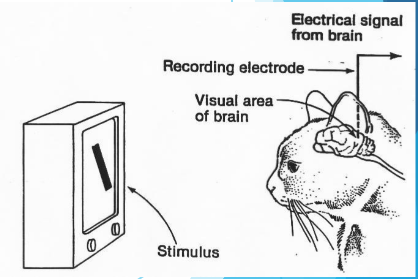
Disclaimer

- A history of the field as captured from multiple sources (including Szeliski's book and other sources credited on each page)
- A slightly biased history, as relevant to the topics we cover in this course. There is more to history in related topics (e.g. geometry-based vision, physics-based vision, image/video processing and compression, graphics, computational photography) not covered herein.
- A slight predisposition to work based on images, more than videos.

Early History[1]



- David Hubel and Torsten Wiesel publish their work "*Receptive fields of single neurons in the cat's striate cortex*"
- Placed electrodes into the primary visual cortex area of an anesthetized cat's brain
- Showed that simple and complex neurons exist, and that visual processing starts with simple structures such as oriented edges



[1]Credit: Rostyslav Demush, medium.com

Early History[2]

1959 1963 1966 1971 '73 1979-82

- World's first digital image: Russell Kirsch and his colleagues develop an apparatus to transform images into number grids
- Image of Russell's infant son: grainy 5cm by 5cm photo, 30,976 pixels (176×176 array)
- Now stored in Portland Art Museum

[2]Credit: Rostyslav Demush, medium.com

Early History[3]

1959 **1963** 1966 1971 '73 1979-82

⇓

- Lawrence Roberts' PhD thesis: *"Machine Perception Of Three-Dimensional Solids"*
- Discussed extracting 3D information about solid objects from 2D photographs of line drawings
- Discussed issues such as camera transformations, perspective effects, and the rules and assumptions of depth perception

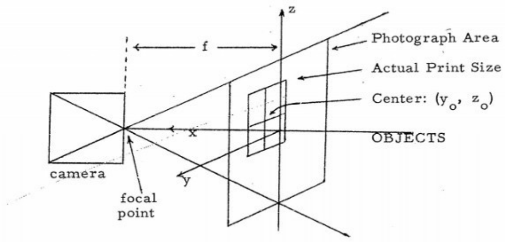
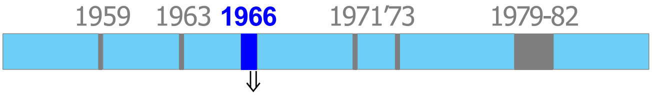


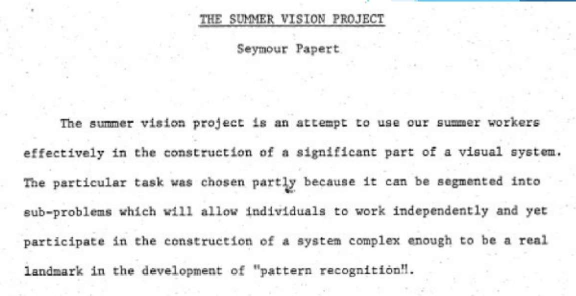
Figure 1: Camera Transformation

[3]Credit: Rostyslav Demush, medium.com

Early History[4]



- Seymour Papert (with Gerald Sussman) from MIT launched the *Summer Vision Project*
- Aimed to develop a platform to automatically segment background/foreground and extract non-overlapping objects from real-world images



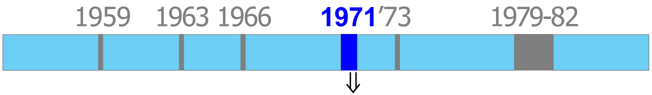
THE SUMMER VISION PROJECT

Seymour Papert

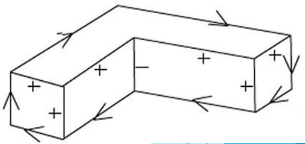
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

[4]Credit: Rostyslav Demush, medium.com

Early History[5]

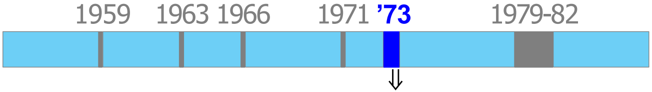


- Discern a shape in a line drawing by labeling lines as convex, concave, and occluded
  - David Huffman et al, *Impossible objects as nonsense sentences*, Machine Intelligence, 8:475-492, 1971
  - Max Clowes et al, *On seeing things*, Artificial Intelligence, 2:79-116, 1971

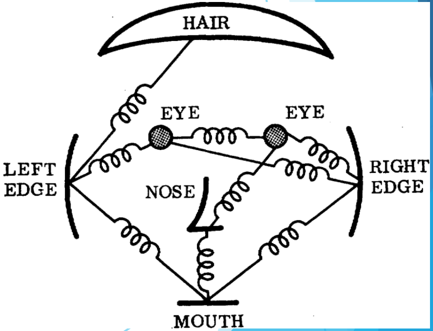


[5]Credit: Rostyslav Demush, medium.com

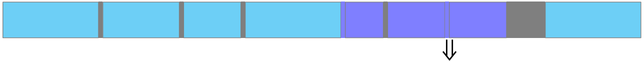
Early History



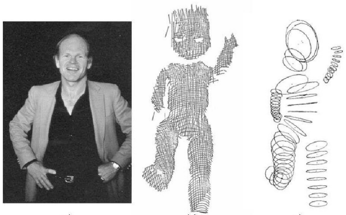
- *Pictorial Structures* model by Fischler and Elschlager
- Given a visual object's description, find the object in a photograph
- Part of the solution is specification of a descriptive scheme, and a metric on which to base the decision of "goodness" of matching or detection

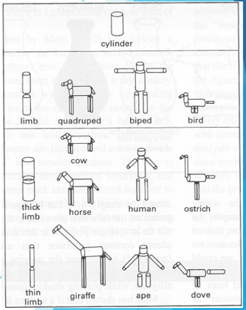


Early History

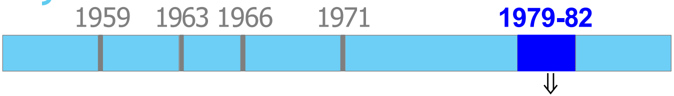


- Object recognition through shape understanding
    - Binford 1971, Generalized Cylinders
    - Marr and Nishihara 1978, Skeletons and Cylinders
- MIT's Artificial Intelligence Lab offers a "Machine Vision" course

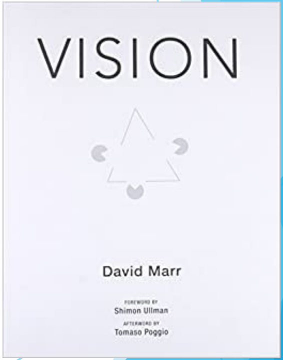




Early History[6]

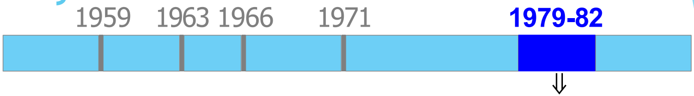


- David Marr, *Vision: A computational investigation into the human representation and processing of visual information*, 1982
- Established that vision is hierarchical
- Introduced a framework where low-level algorithms that detect edges, curves, corners, etc., are used to get high-level understanding of visual data



[6]Credit: Rostyslav Demush, medium.com

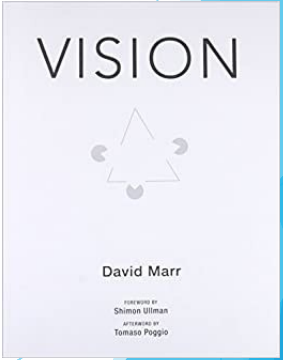
Early History[7]



Marr's Representational Framework

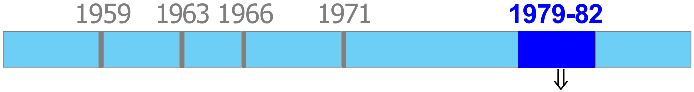
A primal sketch of an image, where edges, bars, boundaries etc., are represented

- A $2\frac{1}{2}$-D sketch representation where surfaces, information about depth, and discontinuities on an image are pieced together
- A 3D model that is hierarchically organized in terms of surface and volumetric primitives

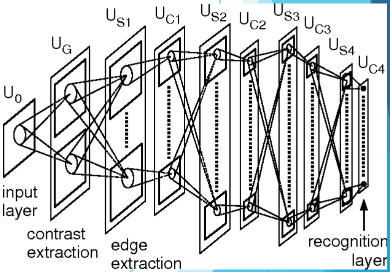


[7]Credit: Rostyslav Demush, medium.com

Early History[8]



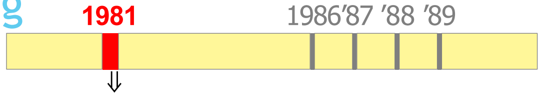
- Kunihiko Fukushima’ *Neocognitron*, a self-organizing artificial network of simple and complex cells to recognize patterns, unaffected by position shifts
- The original ConvNet!
- Included convolutional layers with weight vectors (known as filters)



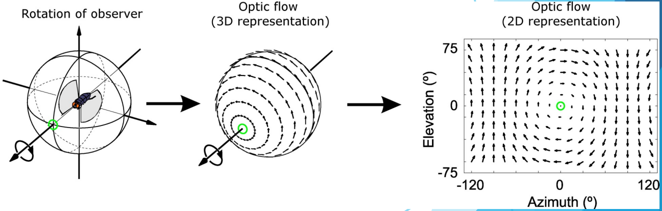
[8]Credit: Rostyslav Demush, medium.com

Part 2

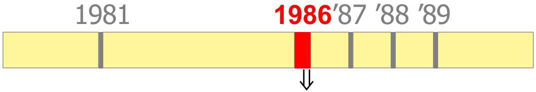
Towards Algorithms and Practice: Low-level Understanding



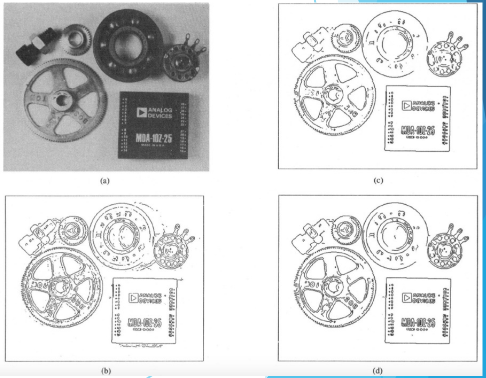
- **Optical Flow:** Horn and Schunck develop method to estimate the direction and speed of a moving object across two images
- Flow is formulated as a global energy functional which is minimized



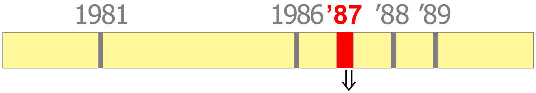
Towards Algorithms and Practice: Low-level Understanding



- **Canny Edge Detector:** Multi-stage edge detection operator, with a computational theory of edge detection
- Used calculus of variations to find the function that optimizes a given functional
- Well-defined method, simple to implement, became very popular for edge detection

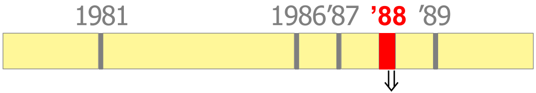


Towards Algorithms and Practice: Low-level Understanding

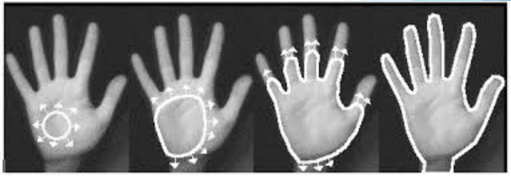


- **Recognition by Components Theory:** Proposed by Biederman
- Bottom-up process to explain object recognition
- Object's component parts: *geons*, based on basic 3-dimensional shapes (cylinders, cones, etc.) assembled to form the object

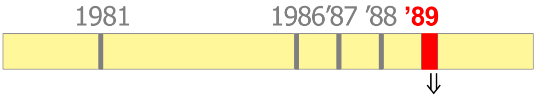
Towards Algorithms and Practice: Low-level Understanding



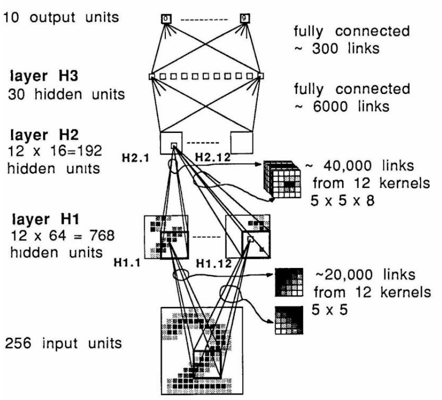
- **Snakes** or **active contour models** delineate an object outline from a possibly noisy 2D image
- Widely used in applications like object tracking, shape recognition, segmentation, edge detection and stereo matching



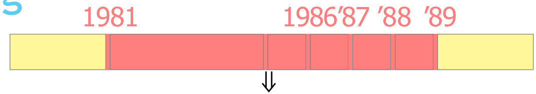
Towards Algorithms and Practice: Low-level Understanding



- **Backprop for CNNs** arrives…
- Applied to handwritten digit recognition provided by USPS



Towards Algorithms and Practice: Low-level Understanding



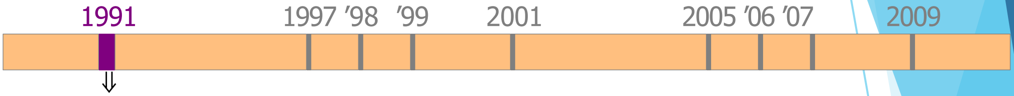
- Image Pyramids[9]
- Scale-space Processing[10]
- Wavelets[11]
- Shape-from-X[12]
- Variational Optimization methods[13]
- Markov Random Fields[14]

---

[9] Burt and Adelson, 1983
[10] Witkin, 1984
[11] Mallat, 1989
[12] Pentland, 1984; Blake et al, 1985
[13] Poggio et al, 1985
[14] Geman and Geman, 1985

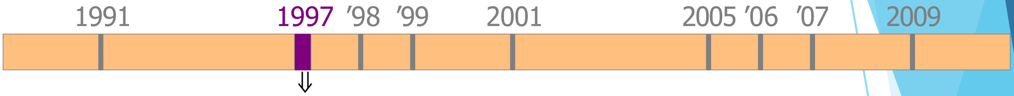
Part 2

Towards Algorithms and Practice: Next-level Understanding



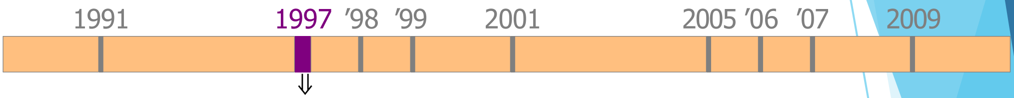
- **Eigenfaces for face recognition** (Turk & Pentland, 1991)

Towards Algorithms and Practice: Next-level Understanding



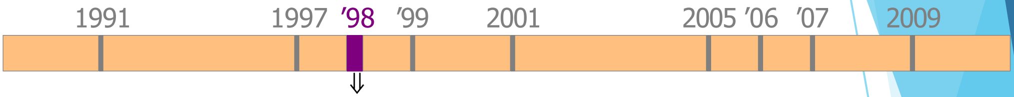
- **Eigenfaces for face recognition** (Turk & Pentland, 1991)
- **Computational theories of object recognition** (Edelman, 1997)

Towards Algorithms and Practice: Next-level Understanding



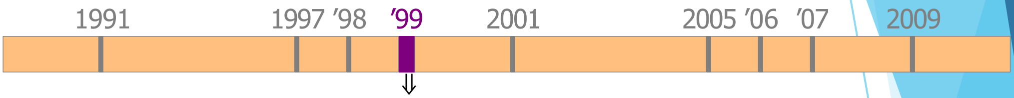
- **Eigenfaces for face recognition** (Turk & Pentland, 1991)
- **Computational theories of object recognition** (Edelman, 1997)
- **Perceptual grouping, Normalized cuts** (Shi & Malik, 1997)

Towards Algorithms and Practice: Next-level Understanding



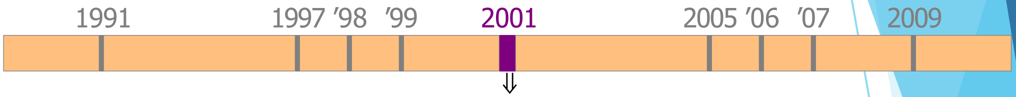
- **Eigenfaces for face recognition** (Turk & Pentland, 1991)
- **Computational theories of object recognition** (Edelman, 1997)
- **Perceptual grouping, Normalized cuts** (Shi & Malik, 1997)
- **Particle filters, Mean shift** for tracking (Liu & Chen, 1998)(Cheng, 1998)

Towards Algorithms and Practice: Next-level Understanding



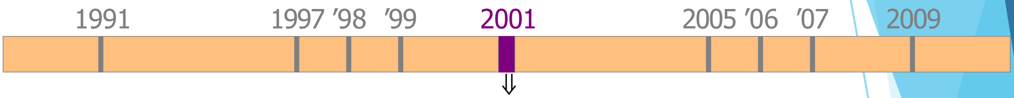
- **Eigenfaces for face recognition** (Turk & Pentland, 1991)
- **Computational theories of object recognition** (Edelman, 1997)
- **Perceptual grouping, Normalized cuts** (Shi & Malik, 1997)
- **Particle filters, Mean shift** for tracking (Liu & Chen, 1998)(Cheng, 1998)
- **SIFT** (Lowe, 1999) (Lowe, 2004)

Towards Algorithms and Practice: Next-level Understanding



- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997) Perceptual grouping, Normalized cuts (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

Towards Algorithms and Practice: Next-level Understanding



- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997) Perceptual grouping, Normalized cuts (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

Conditional Random Fields (Lafferty et al, 2001)

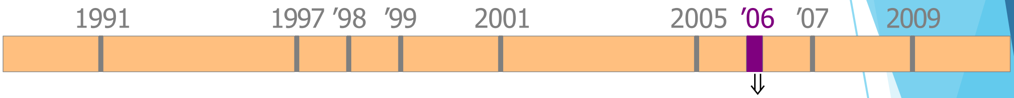
Towards Algorithms and Practice: Next-level Understanding

1991 1997 '98 '99 2001 2005 '06 '07 2009

- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997) Perceptual grouping, Normalized cuts (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

- **Conditional Random Fields** (Lafferty et al, 2001)
- **Pictorial structures** revisited (Felzenszwalb & Huttenlocher, 2005)

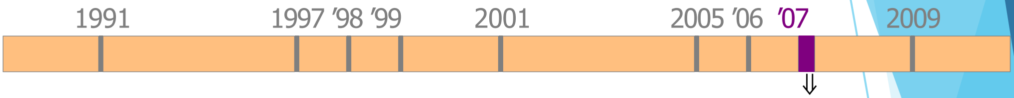
Towards Algorithms and Practice: Next-level Understanding



- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997)
- Perceptual grouping, Normalized cuts
- (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

- **Conditional Random Fields** (Lafferty et al, 2001)
- **Pictorial structures** revisited (Felzenszwalb & Huttenlocher, 2005)
- **PASCAL VOC** arrives; Scene/panorama/location recognition methods grow

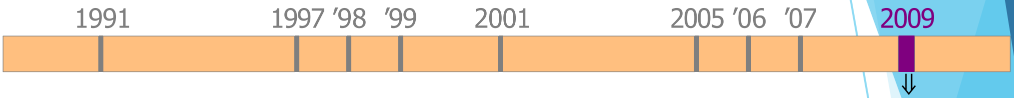
Towards Algorithms and Practice: Next-level Understanding



- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997)
- Perceptual grouping, Normalized cuts
- (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

- **Conditional Random Fields** (Lafferty et al, 2001)
- **Pictorial structures** revisited (Felzenszwalb & Huttenlocher, 2005)
- **PASCAL VOC** arrives; Scene/panorama/location recognition methods grow
- **Constellation models** (Fergus, Perona & Zisserman, 2007)

Towards Algorithms and Practice: Next-level Understanding



- Eigenfaces for face recognition (Turk & Pentland, 1991)
- Computational theories of object recognition (Edelman, 1997)
- Perceptual grouping, Normalized cuts
- (Shi & Malik, 1997)
- Particle filters, Mean shift for tracking (Liu & Chen, 1998)(Cheng, 1998)
- SIFT (Lowe, 1999) (Lowe, 2004)
- Viola-Jones face detection (Viola & Jones, 2001)

- **Conditional Random Fields** (Lafferty et al, 2001)
- **Pictorial structures** revisited (Felzenszwalb & Huttenlocher, 2005)
- **PASCAL VOC** arrives; Scene/panorama/location recognition methods grow
- **Constellation models** (Fergus, Perona & Zisserman, 2007)
- **Deformable part models** (Felzenszwalb et al, 2009)

Part 2

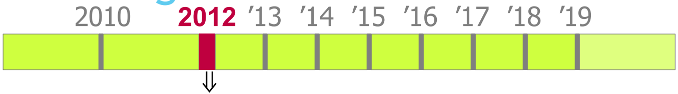
1. Early History: Initial Forays
2. Towards Algorithms and Practice: Low-level Understanding
3. Towards Algorithms and Practice: Next Level of Understanding
4. The Deep Learning Era

The Deep Learning Era

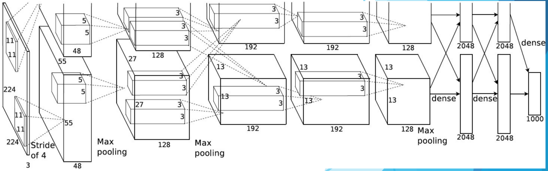
2010 2012 '13 '14 '15 '16 '17 '18 '19

- ImageNet arrives

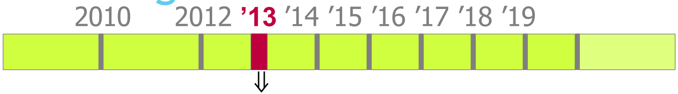
The Deep Learning Era



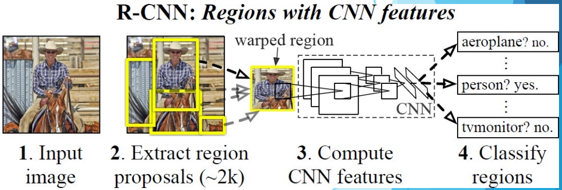
- ImageNet arrives
- AlexNet wins the ImageNet challenge



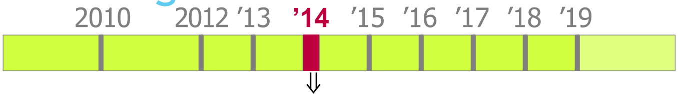
The Deep Learning Era



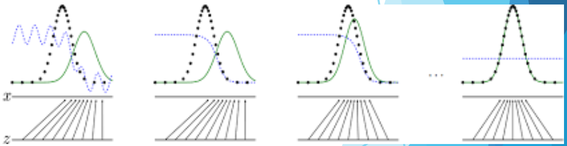
- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins



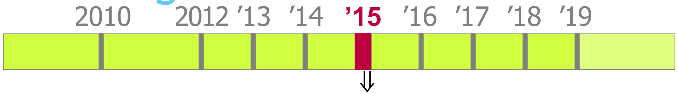
The Deep Learning Era



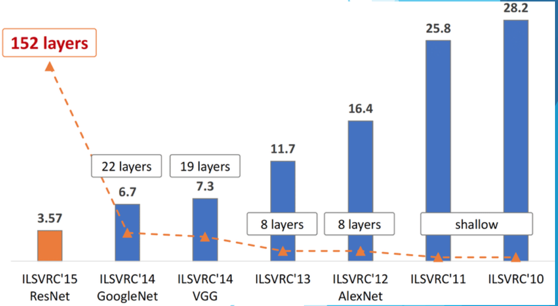
- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins
- InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs



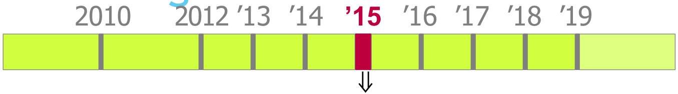
The Deep Learning Era



- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins
- InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs
- ResNet arrives; CNNs match human performance on ImageNet

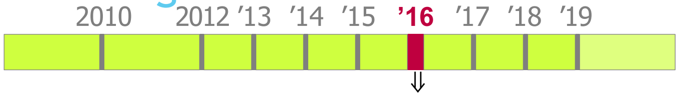


The Deep Learning Era



- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins
- InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs
- ResNet arrives; CNNs match human performance on ImageNet
- FCN, SegNet and U-Net for semantic segmentation; COCO dataset arrives; VQA dataset arrives

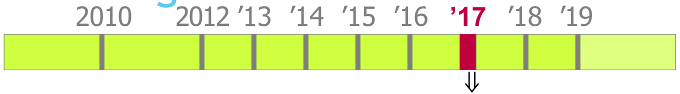
The Deep Learning Era



- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs
- begins InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs
- ResNet arrives; CNNs match human performance on ImageNet

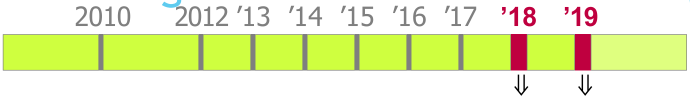
- FCN, SegNet and U-Net for semantic segmentation; COCO dataset arrives; VQA dataset arrives
- YOLO and SSD for object detection; Cityscapes dataset arrives, Visual Genome dataset arrives

The Deep Learning Era



- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins
- InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs
- ResNet arrives; CNNs match human performance on ImageNet
- FCN, SegNet and U-Net for semantic segmentation; COCO dataset arrives; VQA dataset arrives
- YOLO and SSD for object detection; Cityscapes dataset arrives, Visual Genome dataset arrives
- Scene graph generation models

The Deep Learning Era



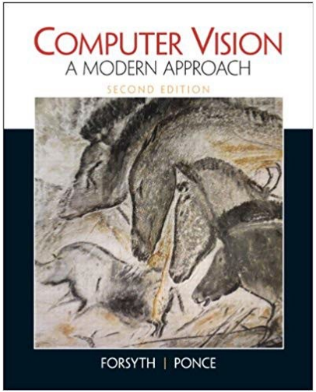
- ImageNet arrives
- AlexNet wins the ImageNet challenge
- A CNN, variant of ZFNet, wins ImageNet challenge; R-CNNs for object detection arrive; Understanding CNNs begins
- InceptionNet, VGG models arrive; Human Pose Estimation CNNs; Deep generative models: GANs, VAEs
- ResNet arrives; CNNs match human performance on ImageNet
- FCN, SegNet and U-Net for semantic segmentation; COCO dataset arrives; VQA dataset arrives
- YOLO and SSD for object detection; Cityscapes dataset arrives, Visual Genome dataset arrives
- Scene graph generation models
- VCR dataset, Panoptic segmentation
- ...

History of Applications[15]

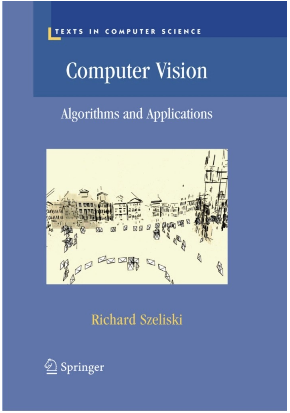
- **1970s:** Optical Character Recognition (OCR)
- **1980s:** Machine vision, Smart cameras
- **1990s:** Machine vision in manufacturing environments, Biometrics, Medical imaging, Recording devices, Video surveillance
- **2000s:** More biometrics, Better medical imaging, Object/Face detection, Autonomous navigation, Google Goggles, Vision on social media
- **2010s:** Everywhere around us

[15]See https://www.phase1vision.com/resources/timeline for a longer historical timeline

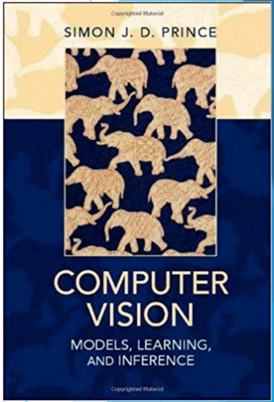
Traditional Computer Vision: References



Book website

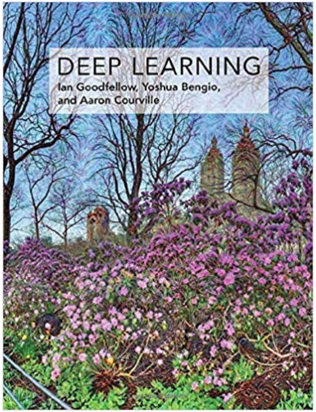


Book website

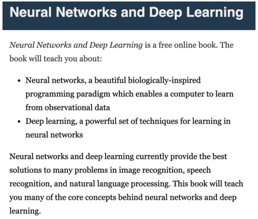


Book website

Resources and References



Book website



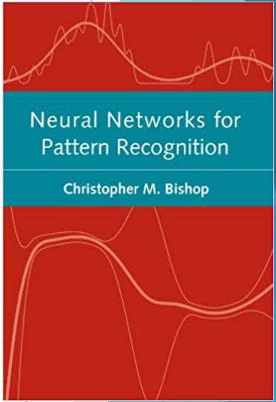
Neural Networks and Deep Learning

Neural Networks and Deep Learning is a free online book. The book will teach you about:

- Neural networks, a beautiful biologically-inspired programming paradigm which enables a computer to learn from observational data
- Deep learning, a powerful set of techniques for learning in neural networks

Neural networks and deep learning currently provide the best solutions to many problems in image recognition, speech recognition, and natural language processing. This book will teach you many of the core concepts behind neural networks and deep learning.

A nice, short online book by Michael Nielsen



Book website

Want to Learn Other Topics?

Learning-based Vision

Visual Recognition, Detection, Segmentation, Tracking, Retrieval, etc

Geometry-based Vision

Feature-based Alignment, Image Stitching, Epipolar Geometry, Structure from Motion, 3D Reconstruction, etc

Physics-based Vision

Computational Photography, Photometry, Lightfields, Color Spaces, Shape-from-X, Reflection, Refraction, Polarization, Diffraction, Interference, etc

Book Link:
Multiple View Geometry in Computer Vision

Resources and References

Want to Learn Other Topics?

Learning-based Vision	Visual Recognition, Detection, Seg- mentation, Tracking, Retrieval, etc
Geometry-based Vision	Feature-based Alignment, Image Stitch-ing, Epipolar Geometry, Structure from Motion, 3D Reconstruction, etc
Physics-based Vision	Computational Photography, Photometry, Light-fields, Color Spaces, Shape-from-X, Reflection, Refraction, Polarization, Diffraction, Interference, etc

Book Link:
Physics-Based Vision: Principles and Practice

Resources and References

Homework!

- Go through all links on the Applications of Computer Vision slide (Slide 8) - they are interesting views/reads!

References

- Dave Litwiller. “CMOS vs. CCD: Maturing Technologies, Maturing Markets-The factors determining which type of imager delivers better cost performance are becoming more refined.”.In: *Photonics Spectra* 39.8 (2005), pp. 54–61.
- Richard Szeliski. *Computer Vision: Algorithms and Applications*.Texts in Computer Science. London: Springer-Verlag, 2011.
- David Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*.2 edition. Boston: Pearson Education India, 2015.
- VSBytes Team. *DSLR Cameras vs Smartphone - Which of the two cameras is better?*