Deep Learning for Computer Vision

Spring 2019

http://vllab.ee.ntu.edu.tw/dlcv.html (primary)

https://ceiba.ntu.edu.tw/1072CommE5052_ (grade, etc.)

Yu-Chiang Frank Wang 王鈺強, Associate Professor

Dept. Electrical Engineering, National Taiwan University

2019/02/20

深度學習於電腦視覺
Deep Learning for Computer Vision

- Time: Wednesdays 234 (09:10-12:10)
- Location: BL-113
- Website:
  http://vllab.ee.ntu.edu.tw/dlcv.html (primary)
  https://ceiba.ntu.edu.tw/1072CommE5052_ (grade, etc.)
  https://www.facebook.com/groups/303387027038348/ (FB: DLCV Spring 2019)
- Email: ntudlcvta2019@gmail.com
- Required Knowledge & Skills
  - Knowledge of linear algebra, vector calculus, and basic probability **(required)**
  - Programming skills **(preferably Python)**
  - Backgrounds in machine learning **(optional)**

What to Expect from this Course?

- Learning-based computer vision
  - Fundamentals of machine learning
  - Visual representation, synthesis, classification, and beyond
  - Deep learning technologies for visual applications
  - Differ from the course of Computer Vision in Fall 2018 (more geometry-based computer vision)
- Practical experiences
  - Assignments and projects dealing with real-world visual data
  - Final projects possibly supported by industries with prizes

- Lots of work and fast paces, but hopefully helpful with lots of fun!

Disclaimer

- This course is taught for the 2nd time at NTU.
- Syllabus, course policy, and HW/project details might change over time. (DL/CV/AI are fast growing/changing areas.)
- This course will be quite demanding and challenging!

Do I need to have ML backgrounds?

- Not exactly but preferably so.
- Will review/highlight some necessary ideas next week.
  (e.g., Bayesian, linear algebra, dimension reduction, nearest neighbor, clustering, training/testing, cross-validation, SVM, kernels)

Course Information

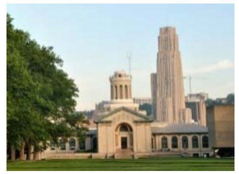
- **加簽原則** (How to sign up if not already in?)
- 講師/助教群介紹 Teaching Team & Office Hours
- 課程大綱與精神
- 成績計算方式

加簽原則

- Capacity
  379
  - 教室容量**170**人 (classroom capacity: 170)
  - 目前初選登記**3~~50~~**人，已選上**100**人
  - 加簽上限以教室容量為準
  - 目前**不能保證**加簽同學運算資源，請自行衡量!!
- Priority
  - 電資學院 & 工學院 (有相關研究需求) > 理學院 > 其他 (EECS ~ Engr. w/ relevant backgrounds > Sci. > others)
  - 博班 > 碩班 > 碩0 > 大四 > 大三 (Based on seniority)
- <u>有意加簽者，請於第二節上課前完成登錄。</u>
  - 登錄網址: **https://goo.gl/2c7kSV**
  - 第二節下課公布加選結果，請準備學生證或相關證件領取授權碼

Teaching Team & Office Hours

- Instructor: Yu-Chiang Frank Wang (王鈺強)
- Research Areas
  - Computer Vision, Machine Learning, Deep Learning, & Artificial Intelligence
- Education
  - PhD, ECE, Carnegie Mellon University, 2009
  - MS, ECE, Carnegie Mellon University, 2004
  - BS, EE, National Taiwan University, 2001

- Contact Info
  - Email: ycwang@ntu.edu.tw
- Office Hour for DLCV Spring 2019
  Preferably after class, appointment by email is suggested.

TAs & Office Hours:

楊福恩
Tue. 1-3pm
BL-527

李元顥
Wed. 1-3pm
BL-527

郭冠軒
Thu. 2-4pm
BL-527

TA Hours: as listed above, or by appointment
TA email: ntudlcvta2019@gmail.com
Location: BL-527 or BL-421

You can also discuss/Q&A with classmates & TAs at Facebook Page:
https://www.facebook.com/groups/303387027038348/
(Search for DLCV Spring 2019)

TAs & Office Hours:

陳尚甫
Fri. 1-3pm
BL-527

吳致緯
Thu. 10am-12pm
BL-421

劉致廷
Wed. 3-5pm
BL-421

TA Hours: as listed above, or by appointment
TA email: ntudlcvta2019@gmail.com
Location: BL-527 or BL-421

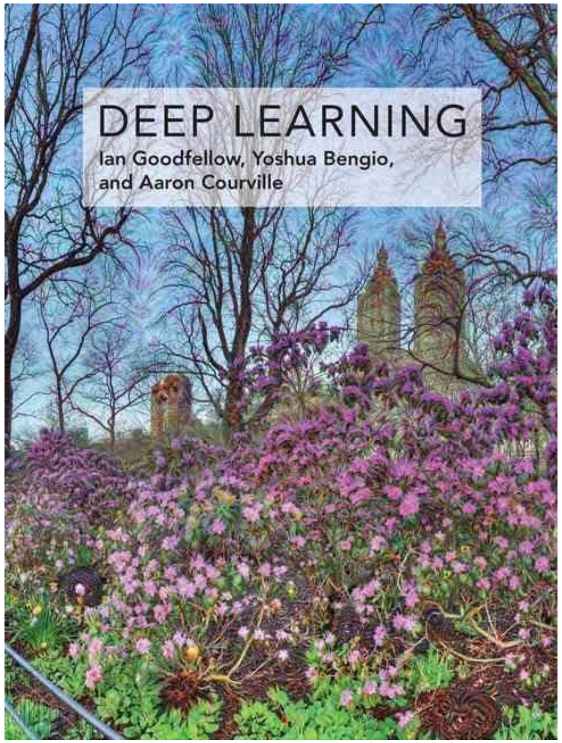
You can also discuss/Q&A with classmates & TAs at Facebook Page:
https://www.facebook.com/groups/303387027038348/
(Search for DLCV Spring 2019)

Tight (yet tentative) Schedule

Week	Date	Topic	Remarks
0	2/20	**Course Logistics + Intro to Computer Vision**	
1	2/27	**Machine Learning 101**	
2	3/06	**Image Representation: From Recognition to Tracking**	HW #1 out
3	3/13	**Intro to Neural Networks + CNN (I)**	
4	3/20	**Intro to Neural Networks + CNN (II)** **Tutorial on Python, GitHub, etc.**	
5	3/27	**Detection & Segmentation**	HW #2 out, HW #1 due
6	4/03	**Spring Break!**	
7	4/10	**Generative Models**	
8	4/17	**Visualization and Understanding NNs**	HW #3 out, HW #2 due
9	4/24	**Transfer Learning for Visual Analysis**	
10	5/01	**Recurrent NNs and Seq-to-Seq Models (I)**	Team Up for Final Projects
11	5/08	**Guest Lecture**	HW #4 out, HW #3 due
12	5/15	**Recurrent NNs and Seq-to-Seq Models (I)**	
13	5/22	**Learning Beyond Images (2D/3D, depth, etc.)**	
14	5/29	**Deep Reinforcement Learning for Visual Apps**	HW #4 due
15	6/05	**Final Project Checkpoint**	
16	6/12 or 6/19	**Final Presentation**	

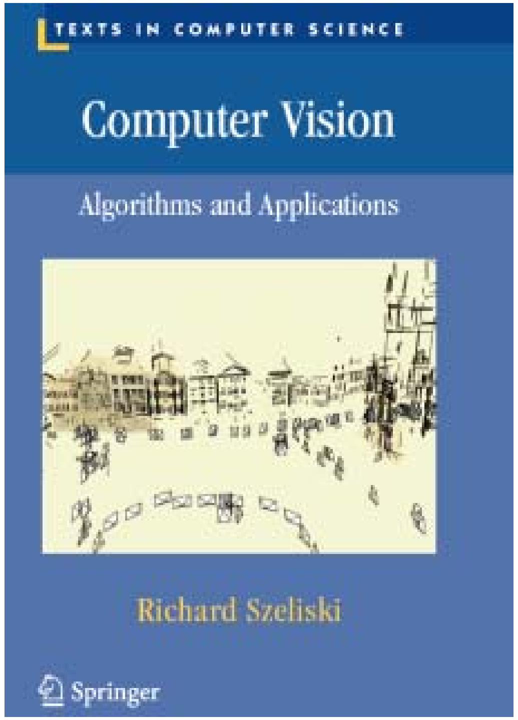
Textbook (Optional)

- Deep Learning, MIT Press
  - Ian Goodfellow, Yoshua Benjio, and Aaron Courville
  - Free online versions available at www.deeplearningbook.org

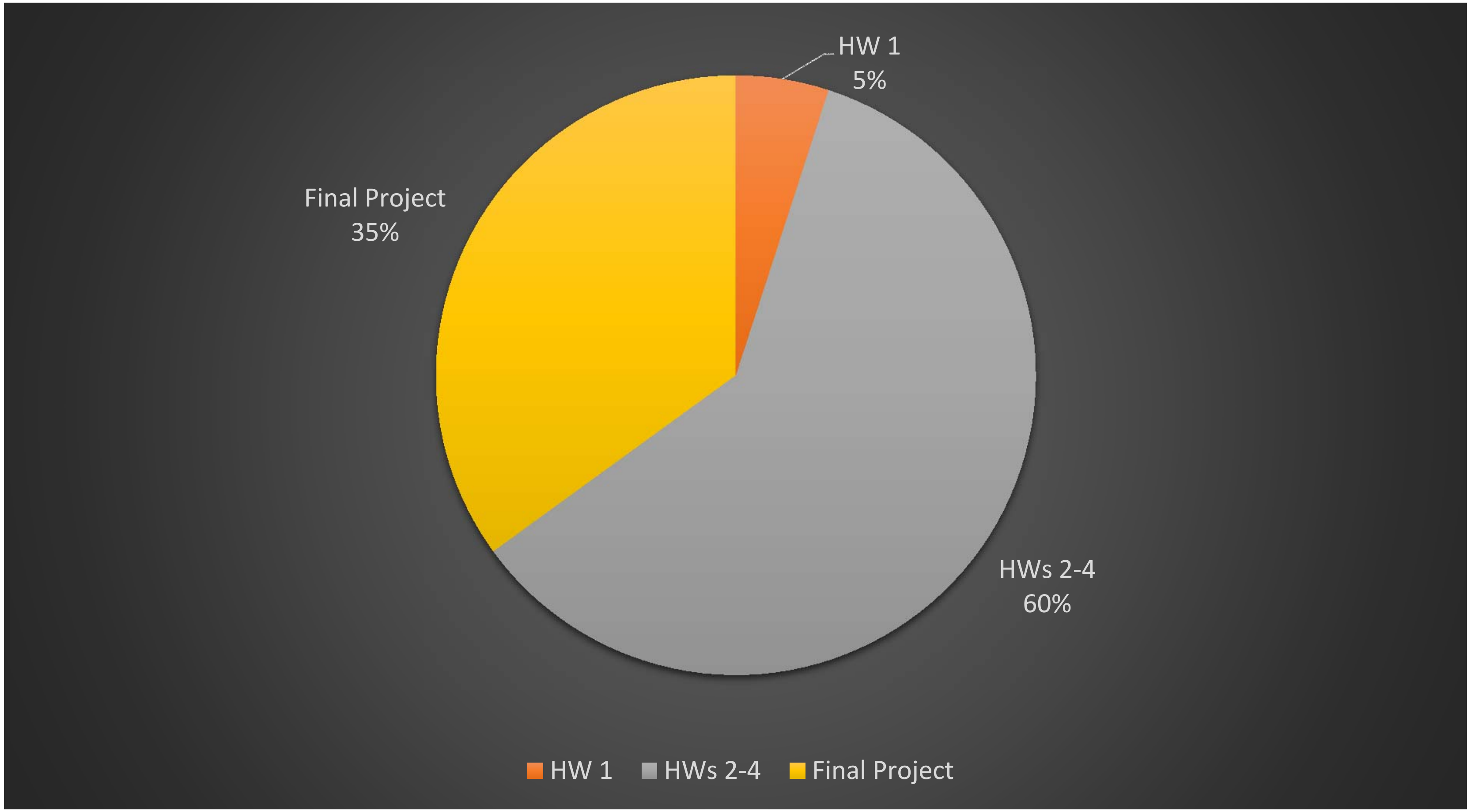


Textbook (Optional)

- Computer Vision: Algorithms and Applications, Springer
  - Richard Szeliski
  - Free online versions available at http://szeliski.org/Book/



About Grading 成績計算方式



Plus up to **5% bonus points**

About Grading (cont'd)

- HW assignments (65%)
  - HW 1 (non-deep HW, 5%), HW 2~4 (20% each)
- Final project + report + poster presentation (35%)
- Participation (bonus up to 5%)
  - Not necessarily 點名
  - Mostly 課堂/facebook參與互動、Q&A等

- *Cash Prize* (up to NTD $???)
  - Selected final projects might be sponsored by industries.
  - Details to be announced.

Final Grade

Letter Grading System	Definition	Grade Points	Conversion Scale
A+	All goals achieved beyond expectation	4.3	90-100
A	All goals achieved	4.0	85-89
A-	All goals achieved, but need some polish	3.7	80-84
B+	Some goals well achieved	3.3	77-79
B	Some goals adequately achieved	3.0	73-76
B- (passing grade for graduate students)	Some goals achieved with minor flaws	2.7	70-72
C+	Minimum goals achieved	2.3	67-69
C	Minimum goals achieved with minor flaws	2.0	63-66
C- (passing grade for undergraduate students)	Minimum goals achieved with major flaws	1.7	60-62
F	Minimum goals not achieved	0	59 and below
X	Not graded due to unexcused absences or other reasons	0	0
W	Withdrawal		
NG	No grade reported		
IP	In progress		
TR	Transfer credit		
EX	Exempted		

http://www.oia.ntu.edu.tw/upload/files/20150616223619.pdf

About Course HWs/Projects

- Computing resource (please be conservative):
  - About USD ~~$120~~ credit per student
  - Supported by Microsoft
  - Details to be announced in March

- About HW late policy
  - 如期中考、專題、社團、約會、~~找不到人抄~~
  - For HWs 2~4, up to **THREE** free late days this semester (e.g., 10min ~ 23hr 59 min count as ONE late day.)
  - After that, a penalty to **30%** per day.
  - **No** late submission for the final project

About Course HWs/Projects

- About Final Project
  - 3~4 (max) people per group
  - Project proposal, mid-term progress report, final report, poster presentation
  - Selected topics come with cash prizes up to NTD $???.
  - Details to be announced after spring break
  - Evaluated by instructor, TAs, and possibly guest judges
  - (Intra/inter-group) peer evaluation will be conducted.
  - Snack/drinks will be provided during final presentation.

Academic Integrity

- Can discuss HW with peers, but DO NOT copy and/or share code
  - **任一次作業抄襲/被抄襲者**，按校規論且**本課程學期成績為F!**
  - This is university policy and not negotiable.
  - Seriously, we gave at least 2 **F**s last year based due to this.
- Do not directly use code from Internet unless you have permissions.
  - If not sure, ask!
  - If so, do specify in your HW/project.
- Do **NOT** use your published work as your final project.
  - However, you are encouraged to extend your previous works.
  - Also, you are encouraged to turn your high-quality projects into publications.

Let’s take a short break!

有意加簽者，請於第二節上課10am前完成登錄。

登錄網址: **https://goo.gl/2c7kSV**

第二節下課公布加選結果，請準備學生證或相關證件領取授權碼

What is Computer Vision?

- **Some Quick Remarks**
  - Give machines *visual perception*
  - Learning for visual data
  - In addition to Machine Learning, computer vision is closely related to Image Processing, Computer Graphics, Computational Photography, etc.

***How many** people are there?*

***What** are people doing?*

***What object** is the guy standing on?*

***Where** is this picture taken?*

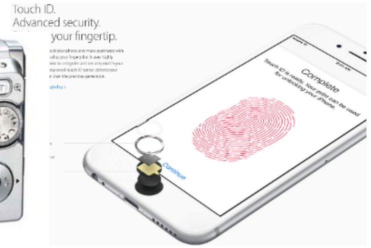
*Why is this picture **funny**?*

Why Computer Vision?

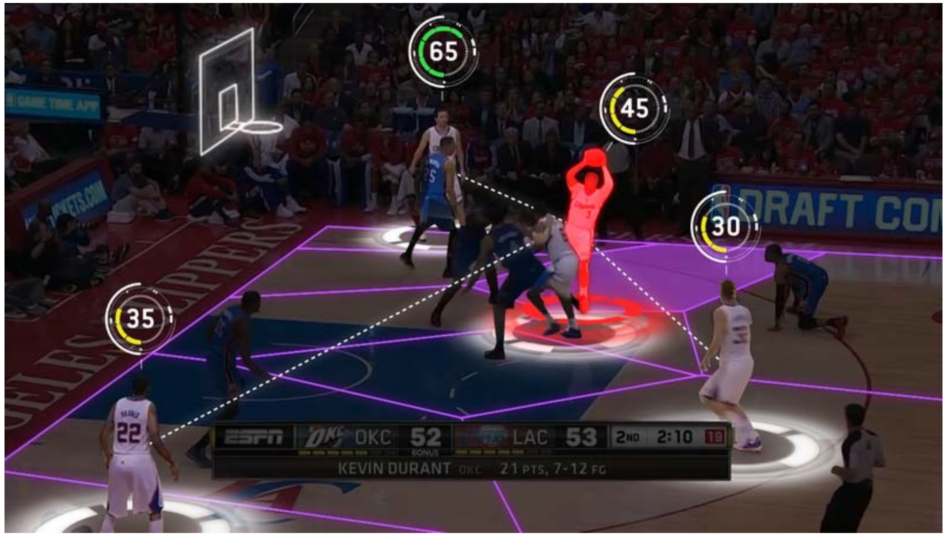
Why & When for Computer Vision?

- Existing CV Applications
  - Biometrics (e.g., face, iris, gait recognition)
  - Optical character recognition (OCR)
  - Sports (tennis, football, basketball, etc.)

  And many more…



Google Translate



Computer Vision in Baseball

Replay Technologies: improving viewer experiences

3D Scanner from a Mobile Camera

MobileFusion

3D from Thousands of Images

[Furukawa et al. CVPR 2010]

VR, AR, and Mixed Reality

Why and When for Computer Vision?

- Exciting AI+CV Applications
  - VR/AR/MR
  - 2D or 3D indoor/outdoor tour from photo collections
  - Style transfer (e.g., *Prisma*: best app of the year)
  - Self-driving car
  - Industrial robots
  - Medical imaging

  And increasingly more than we can imagine!

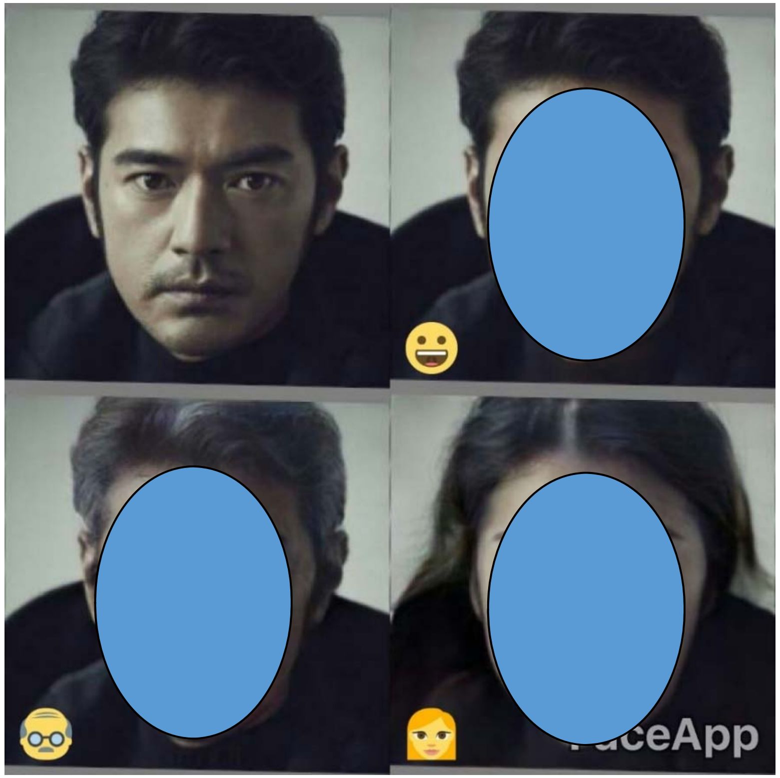
Style Transfer

[EverFilter]

FaceApp

- Deep learning for feature disentanglement
- Beyond putting a smile on your face
- Over 10M downloads

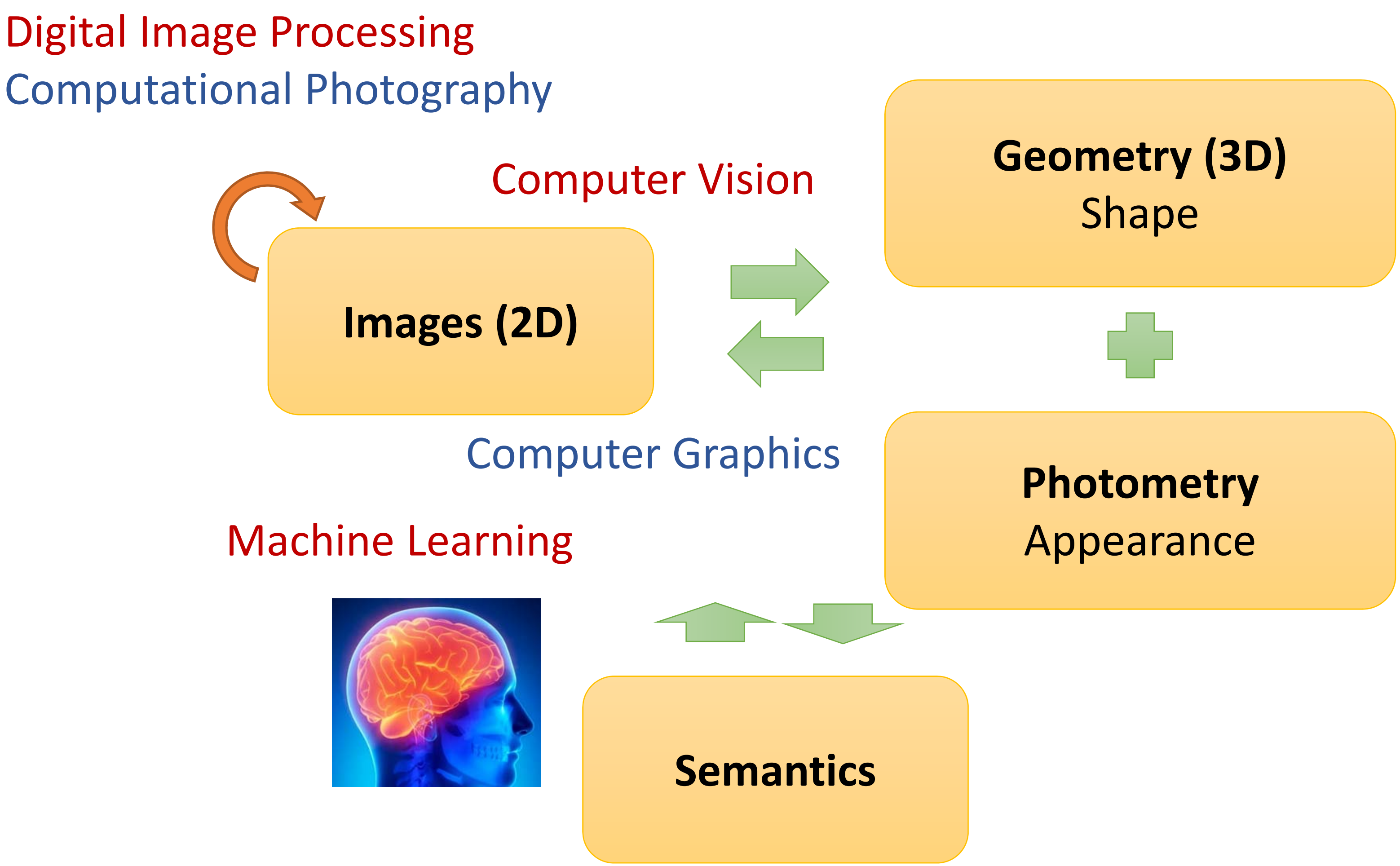
Input



Now, if you are interested in Computer Vision…

Computer Vision and Nearby Fields

- Learning from visual data; learning for visual perception



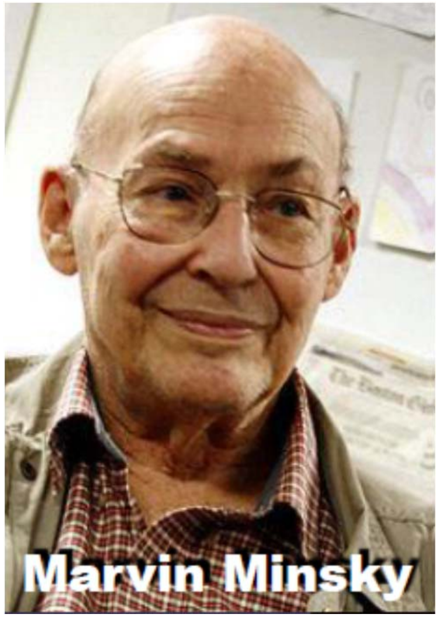
A little bit history of Computer Vision…

Computer Vision:
It can’t be that hard, right?

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Back to 1966…

- In 1966, Minsky hired a first-year undergraduate student, Sussman, and assigned him a problem to solve over the summer: *connect a television camera to a computer and get the machine to describe what it sees*.



Turing Award, 1969

Professor, EE, MIT

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

Half a century later,
we're still working on it.

THE SUMMER VISION PROJECT

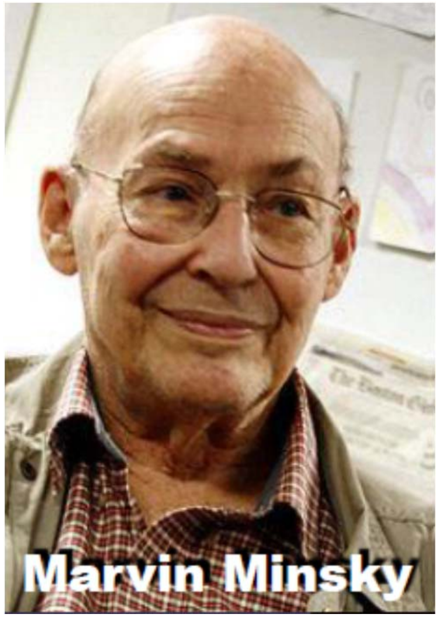
Seymour Papert

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".

Back to 1966...

- “You’ll notice that Sussman never worked in vision again!”

  \- Berthold Horn

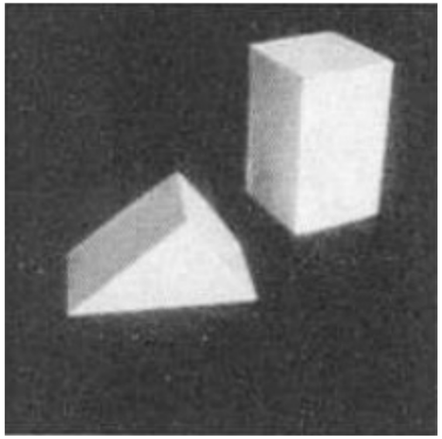


Turing Award, 1969

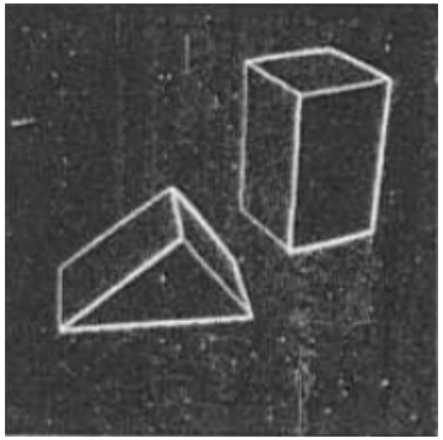
Professor, EE, MIT

1960's: Interpretation of synthetic worlds

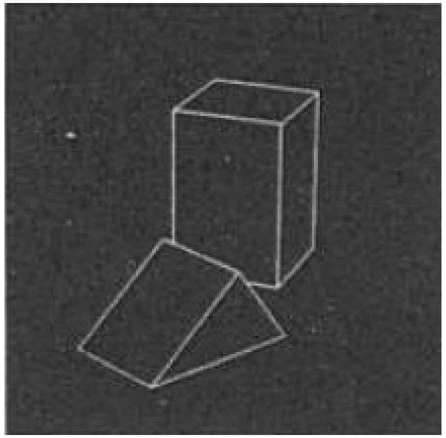
Larry Roberts
"Father of Computer Vision"



Input image



2x2 gradient operator

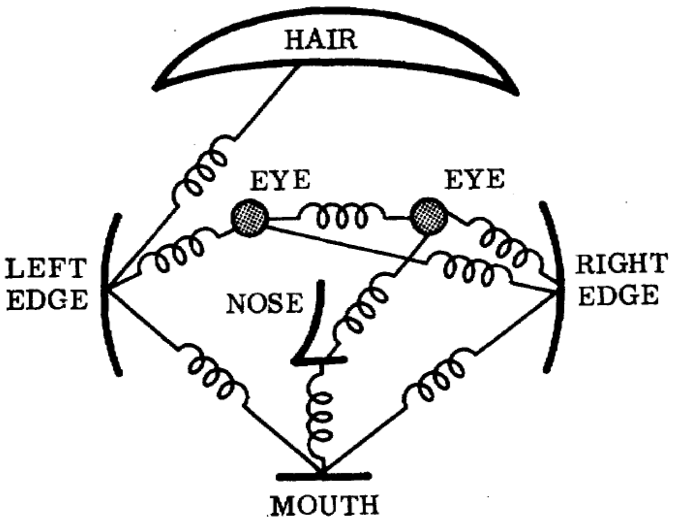


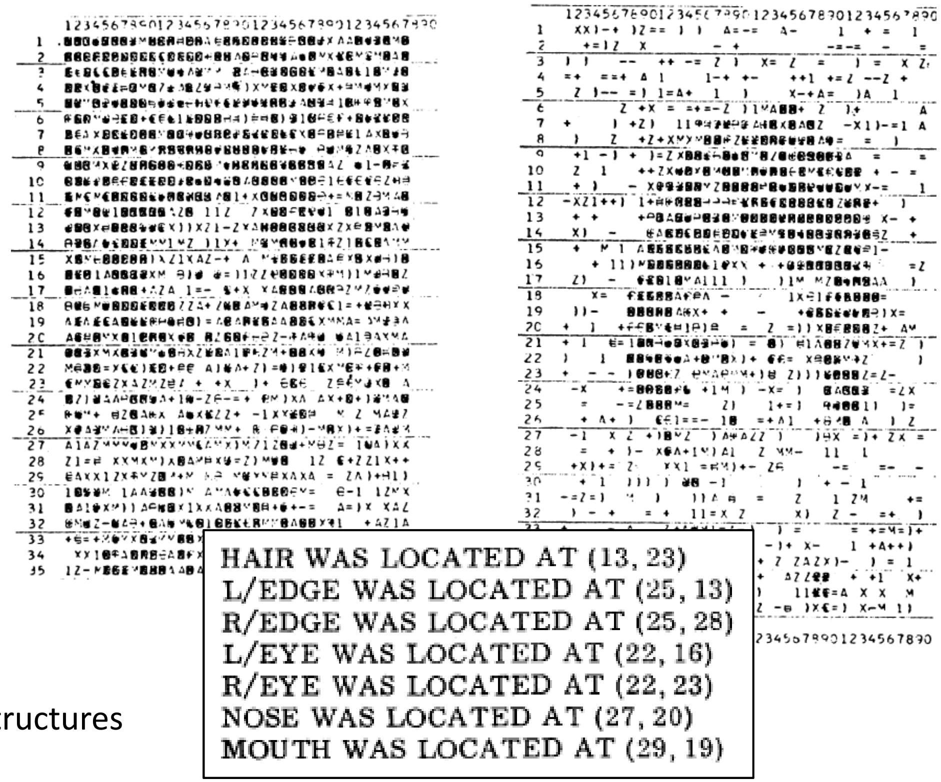
computed 3D model
rendered from new viewpoint

Larry Roberts PhD Thesis, MIT, 1963,
Machine Perception of Three-Dimensional Solids

Slide credit: Steve Seitz

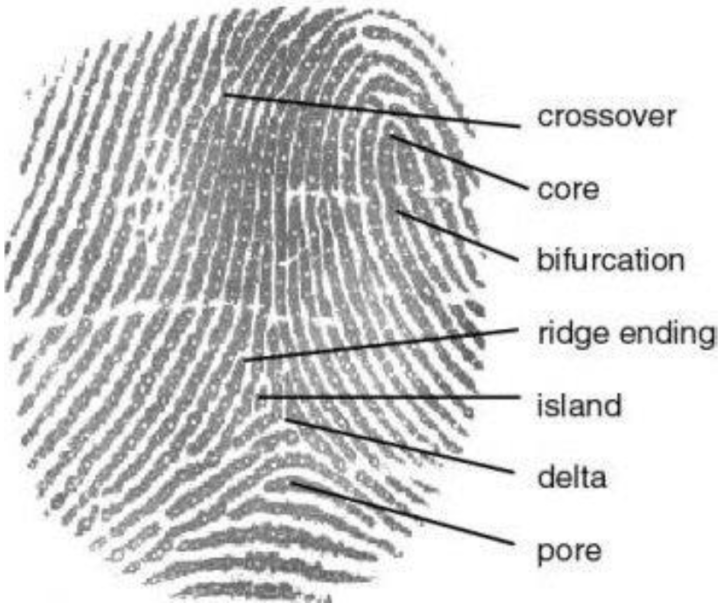
1970’s: Some progress on interpreting selected images (cont’d)

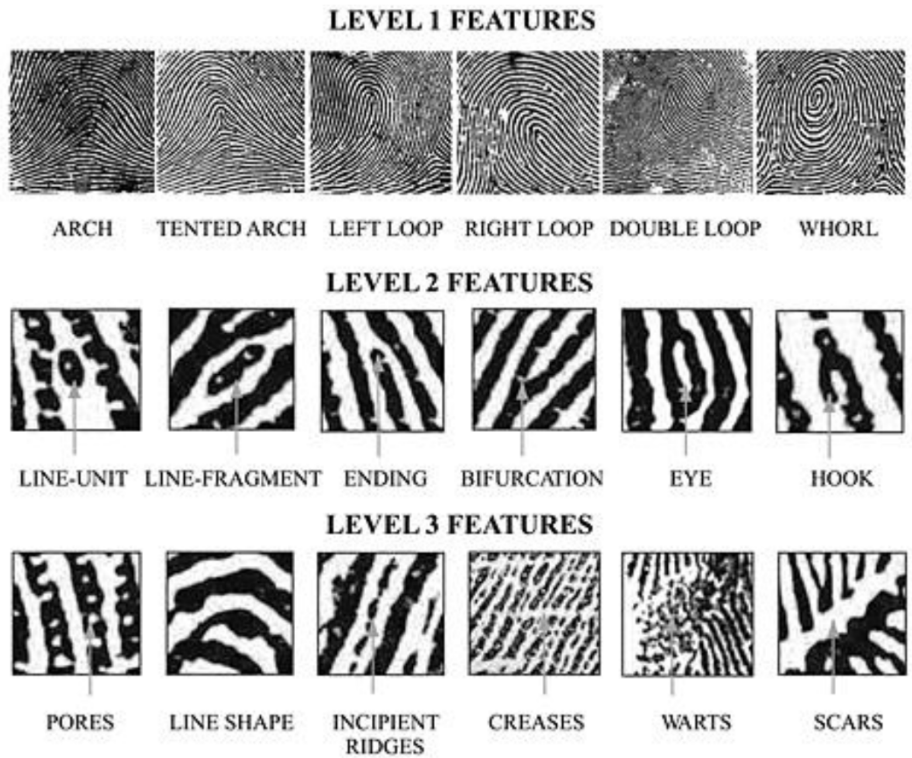




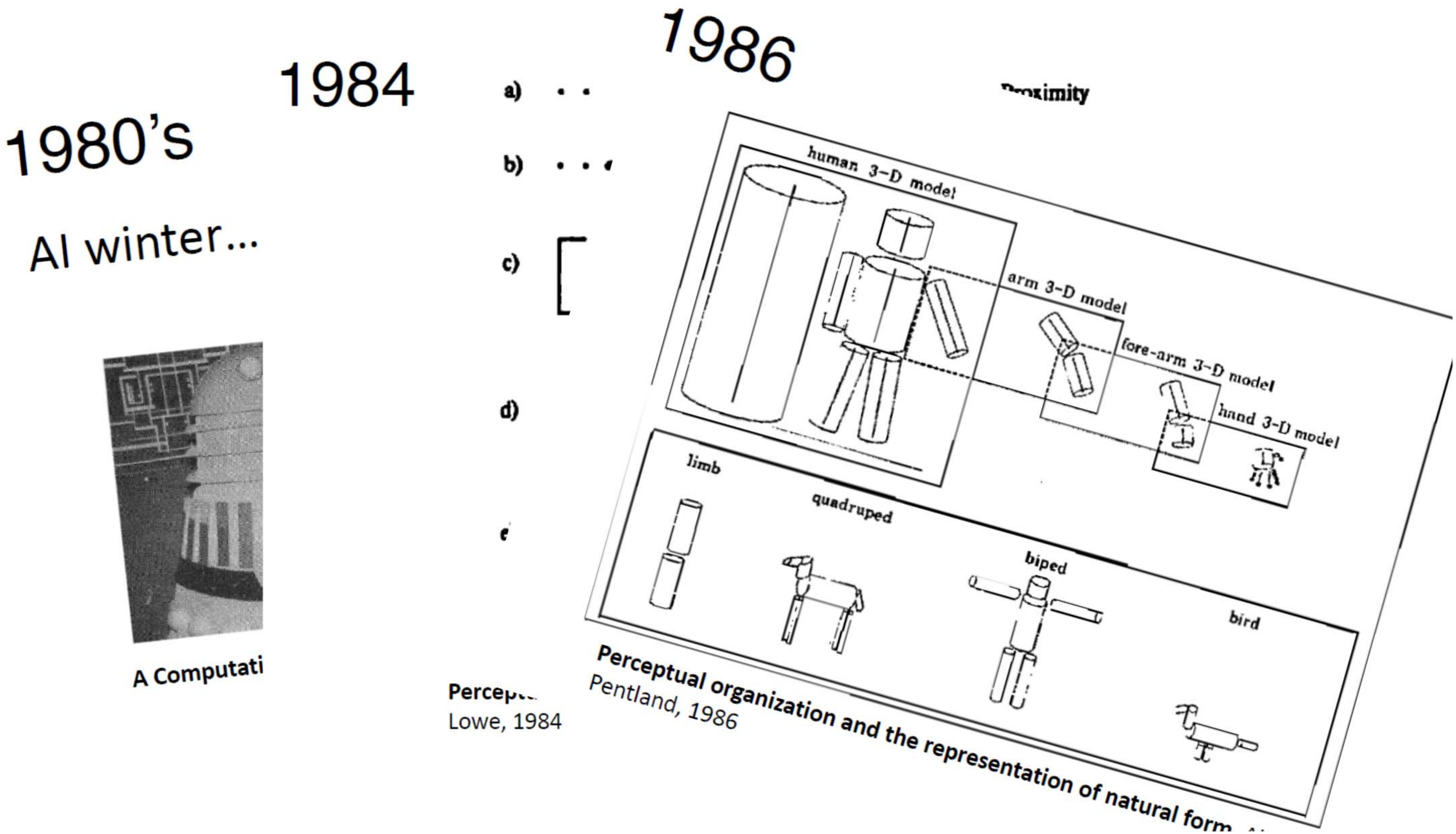
The representation and matching of pictorial structures
Fischler and Elschlager, 1973

1970’s: Some progress on interpreting selected images (cont’d)



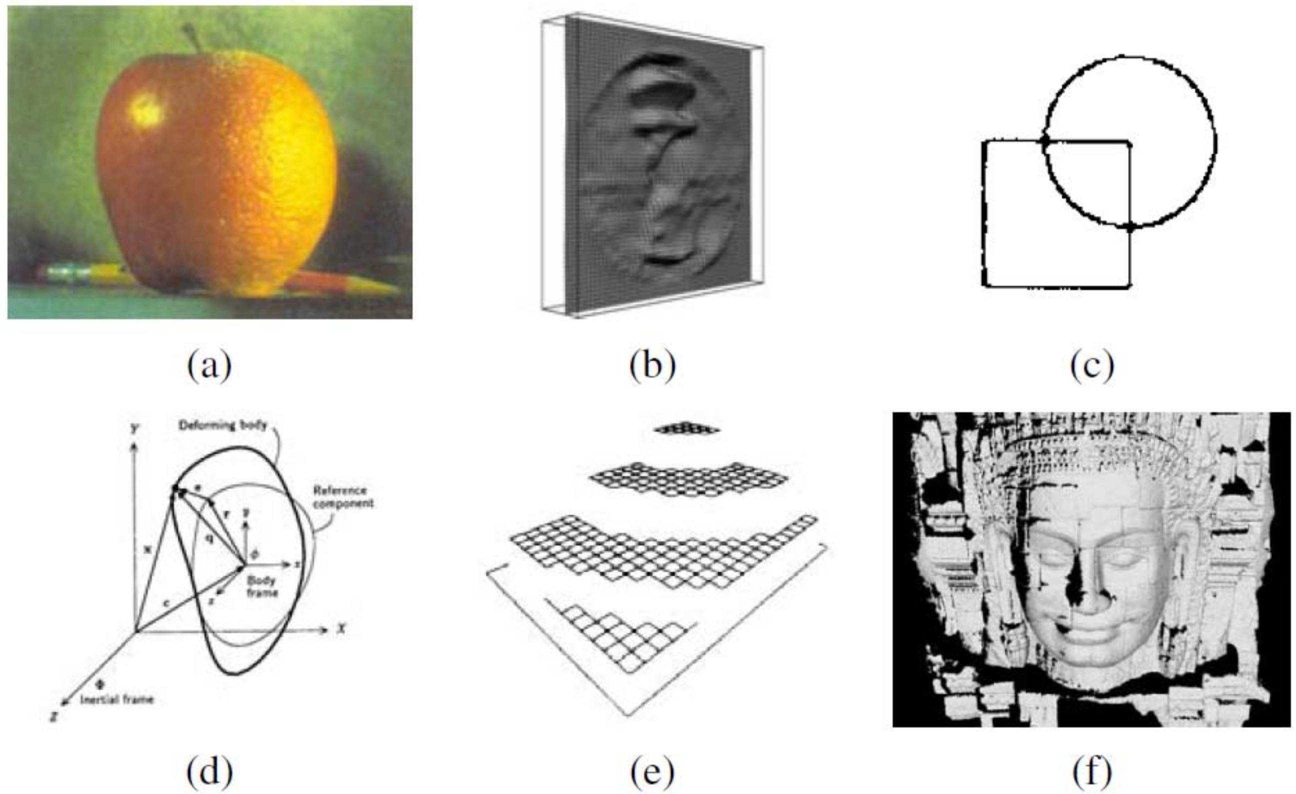


1980's: ANNs come and go (AI winter); shift toward geometry & increased mathematical rigor



1980's: ANNs come and go (AI winter); shift toward geometry & increased mathematical rigor

Blending | Shape from Shading | Edge Detection

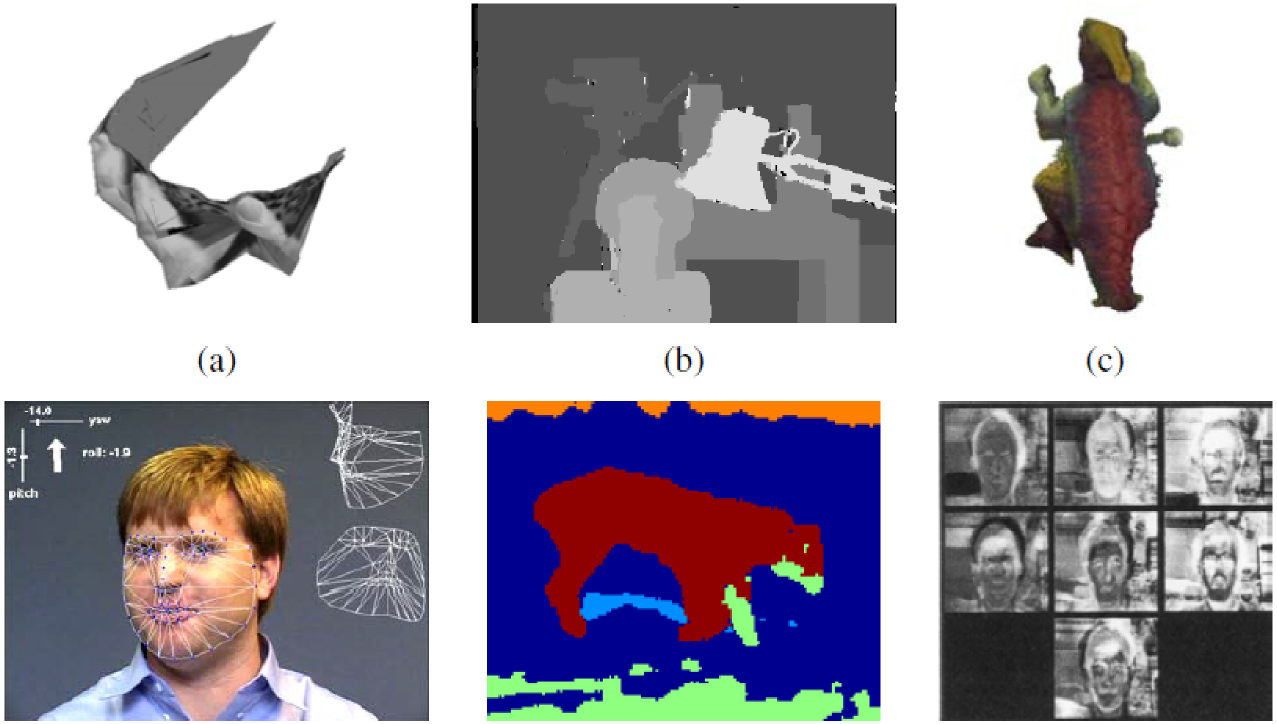


Physics-based methods

Range data acquisition

Image credit: Rick Szeliski

1990's: Structure, Segmentation, and Face Recognition

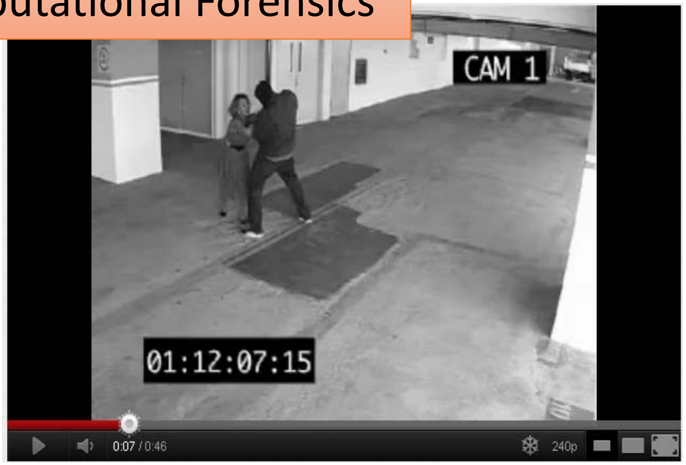


(a) (b) (c) (d) (e) (f)

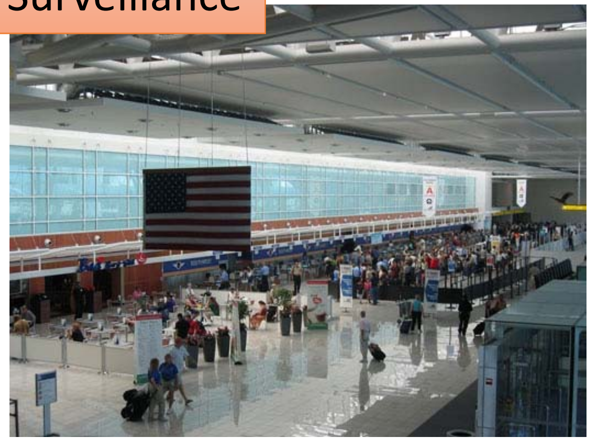
Image credit: Rick Szeliski

1990's: Face recognition (cont'd)

Computational Forensics

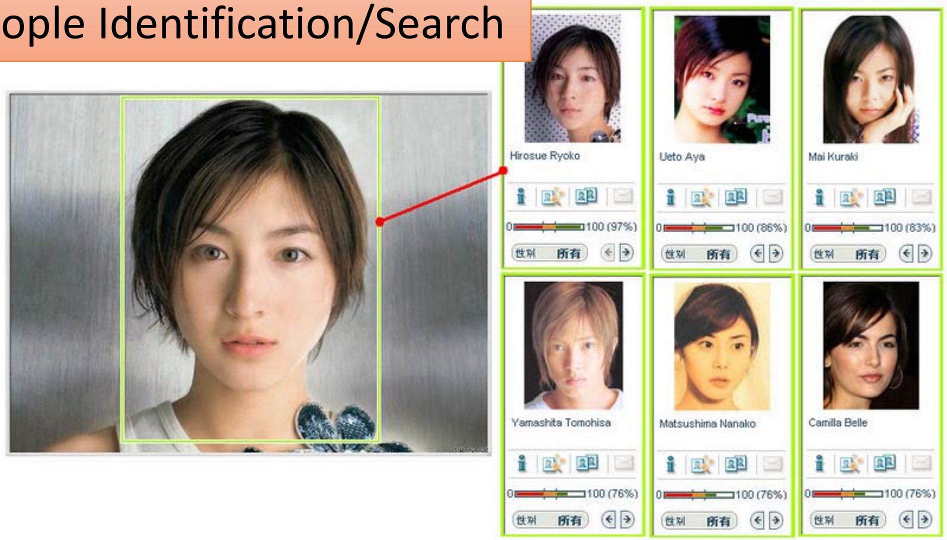


Mass Surveillance



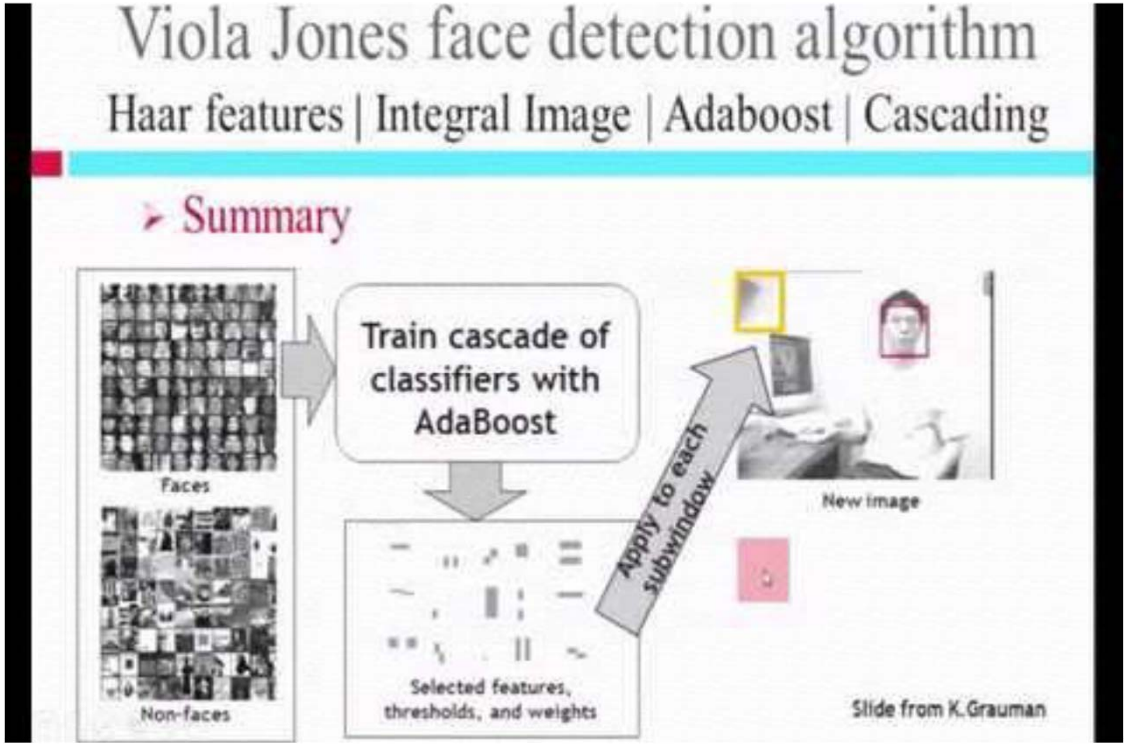
Access Control

People Identification/Search



1990's: Face recognition (cont'd)

- Face verification (1:1) vs. face identification (1:N)
- Face detection by Viola & Jones, 2001



2000's: Broader, Larger and More -
Dataset, Computational Photography, and Video Processing

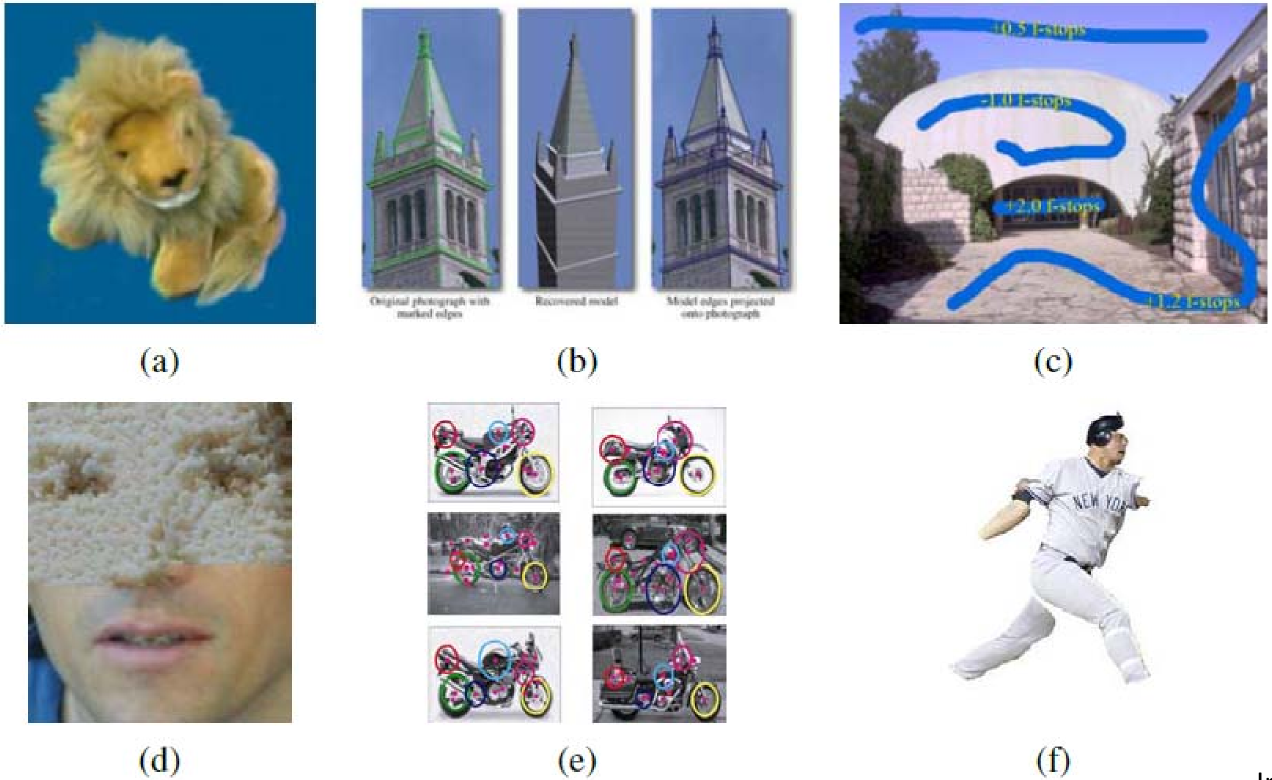
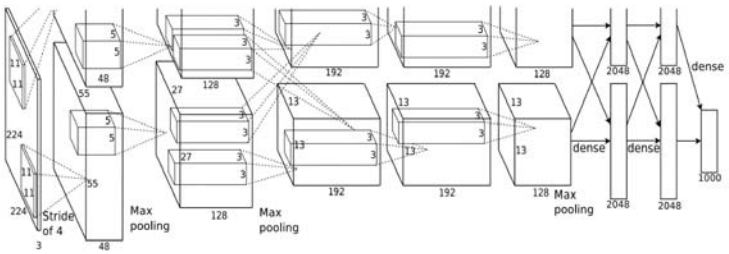
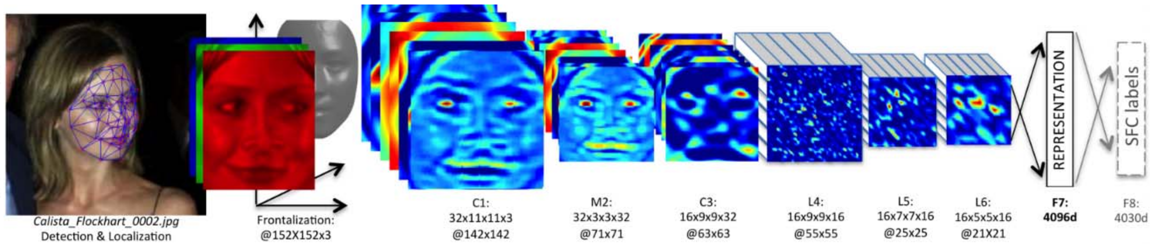


Image credit: Rick Szeliski

2010's:
The ~~Revenge~~ Resurgence of Deep Learning

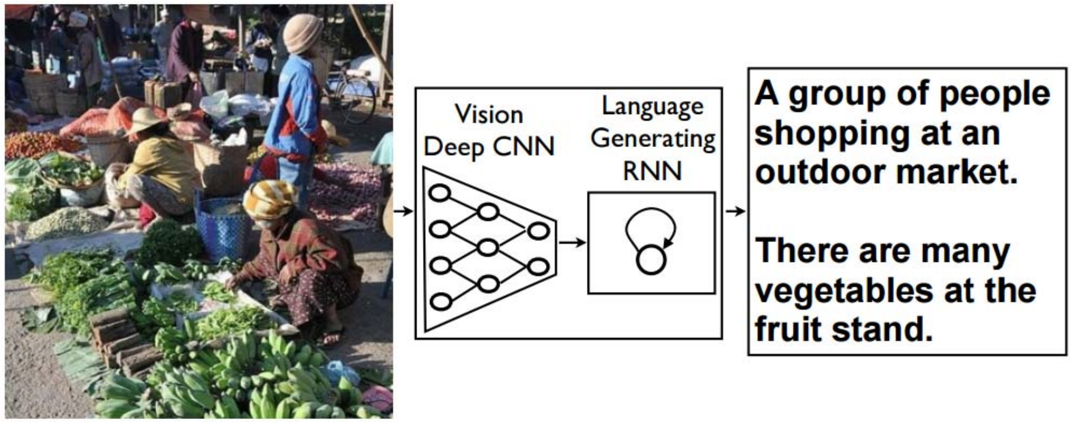


[AlexNet NIPS 2012]



[DeepFace CVPR 2014]

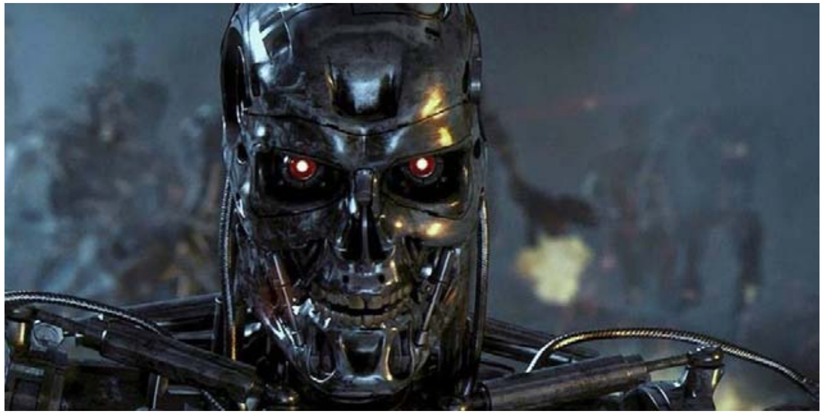
[DeepPose CVPR 2014]

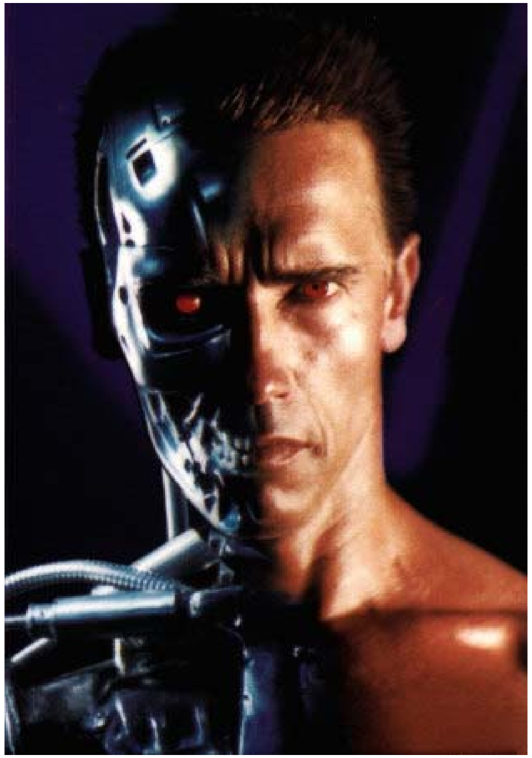


[Show, Attend and Tell ICML 2015]

2020’s: Autonomous vehicles

2030's: The Rising of Robots & Skynet?





Computer Vision with Deep Learning & Artificial Intelligence

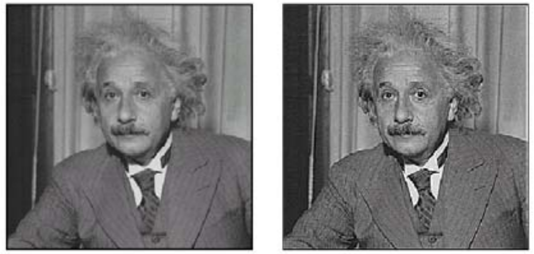
Fundamentals of Computer Vision

- Light
  - What an image records
- Matching
  - How to measure the similarity of two regions
- Alignment
  - How to recover transformation parameters based on matched points
- Grouping
  - What points/regions/lines belong together?
- Geometry
  - How to relate world coordinates and image coordinates
- Recognition
  - What similarities are important?

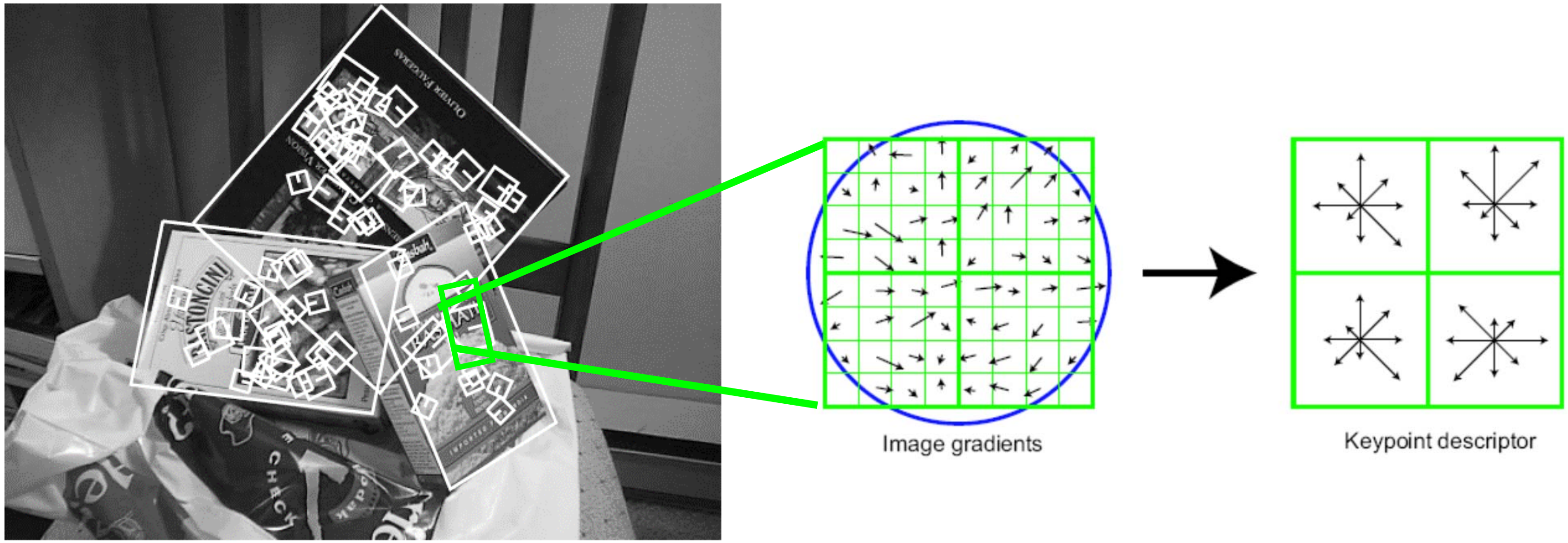
Slide credit: Derek Hoiem

3 Arrows in Computer Vision

- Image Filters
  - Detection of edge, corner, etc.
  - Denoising, smoothing, sharpening, etc.

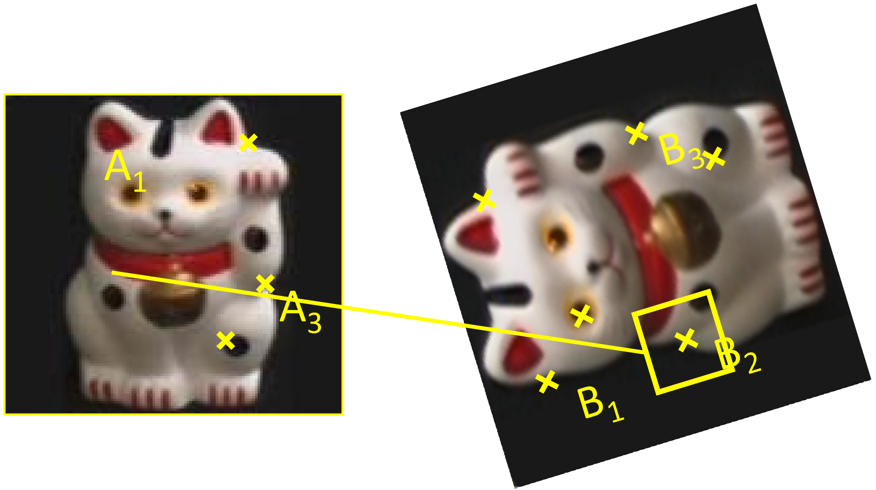


before after



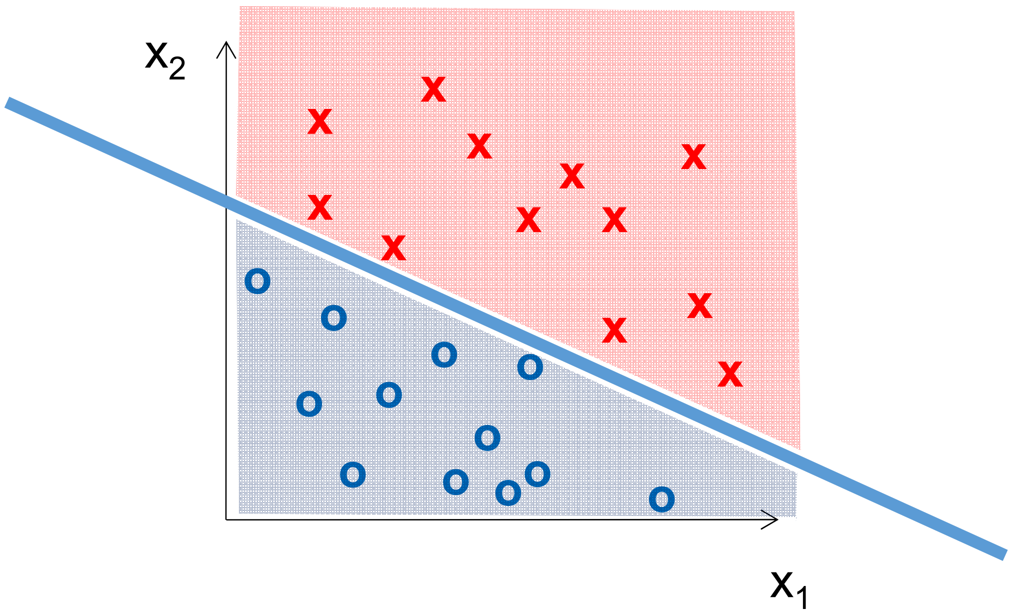
3 Arrows in Computer Vision (cont'd)

- Image Filters
- Feature Descriptors
    - Shift, scale, rotation, etc. invariant
    - Matching & alignment



3 Arrows in Computer Vision (cont'd)

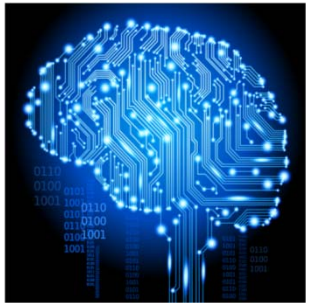
- Image Filters
- Feature Descriptors
- Classifiers
  - Representation, clustering
  - Reasoning, recognition



Deep Learning for Computer Vision

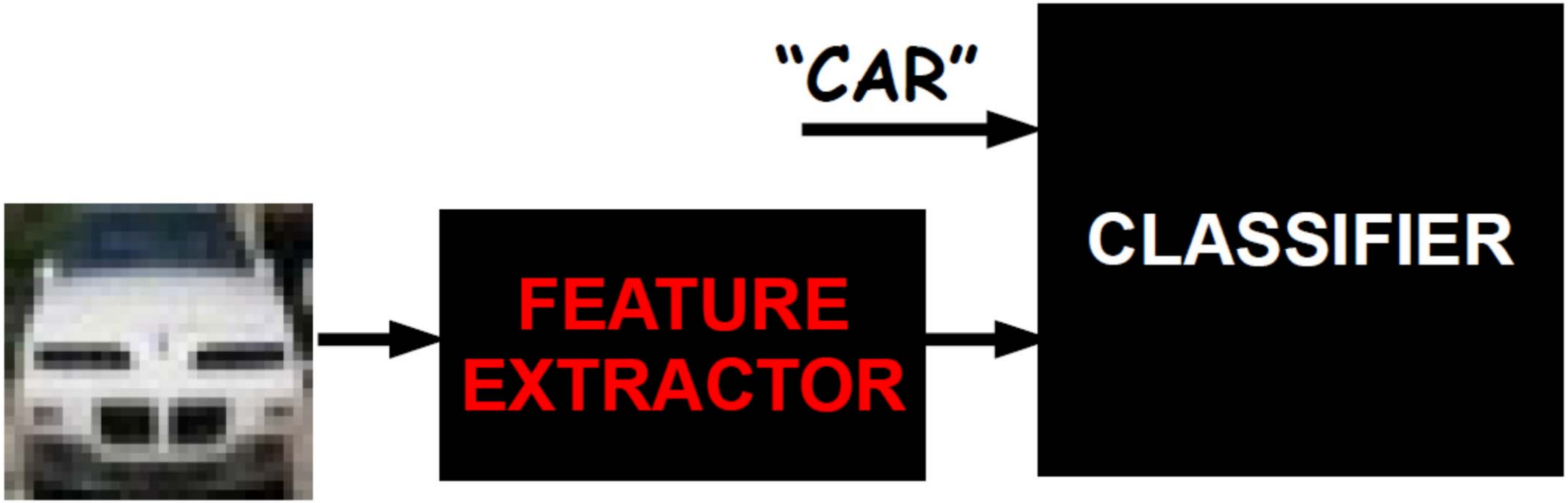
- Image Filters
- Feature Descriptors
- Classifiers

Deep Learning



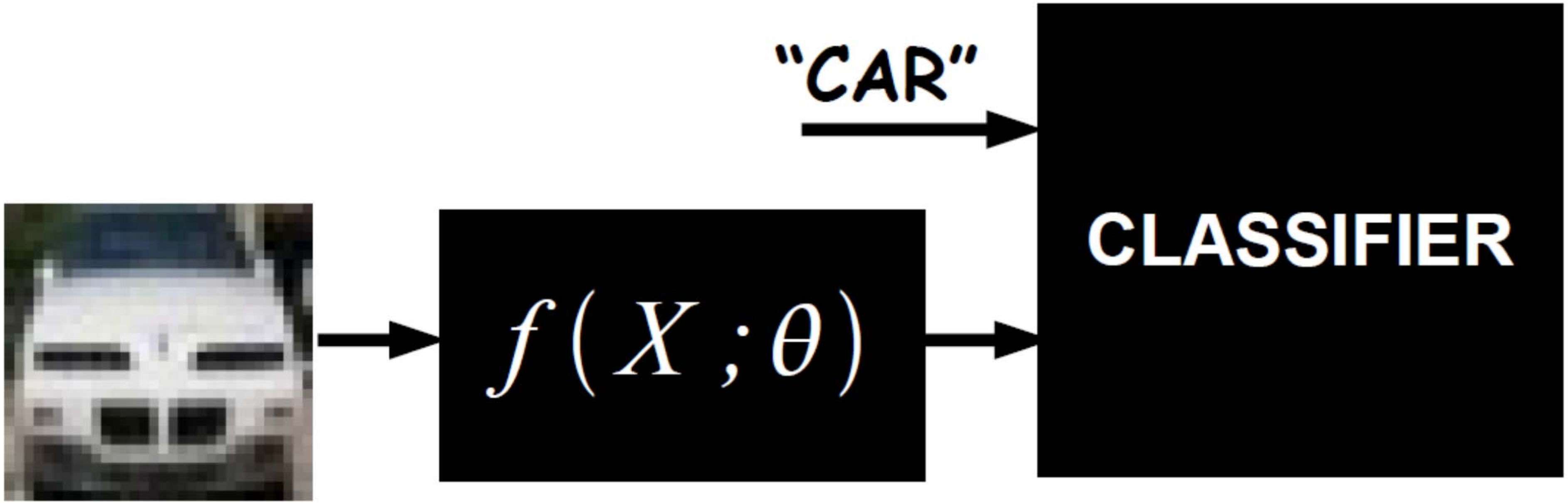
Example:
Building an object recognition system

- Idea: Use data to optimize features for the give task



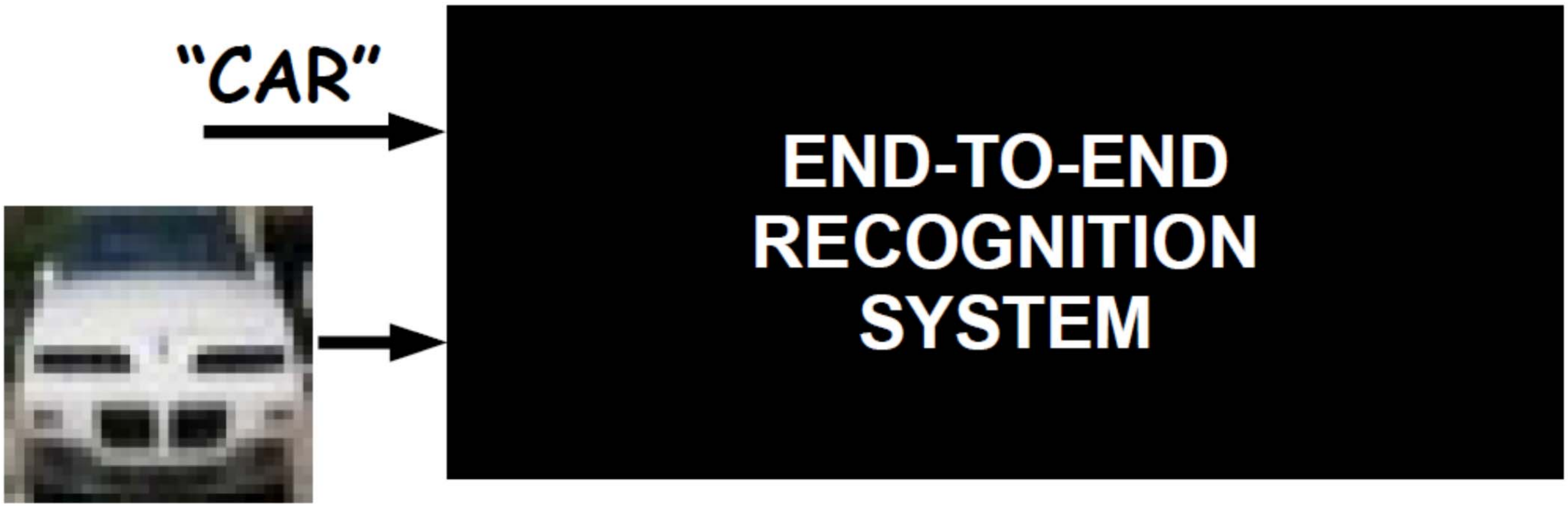
Example: Building an object recognition system (cont'd)

- What we need:
  - Use parameterized functions such that
    Features can be computed and trained efficiently.



Example: Building an object recognition system (cont'd)

- Deep Learning:
  - Everything becomes adaptive.
  - No distinction between feature extractor and classifier.
  - A complex and non-linear system trained from raw data.

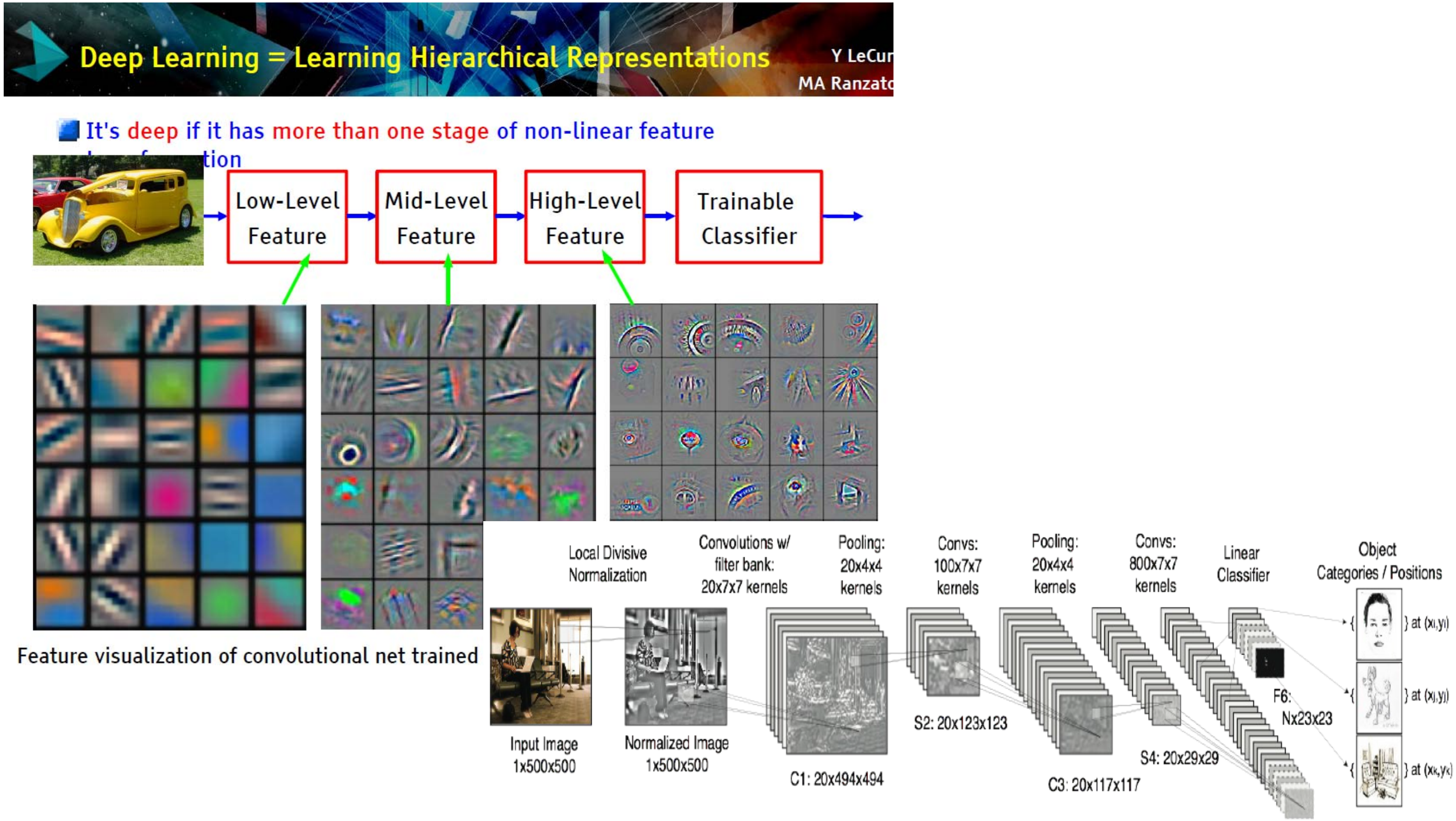


Example:
Building an object recognition system (cont'd)

- Deep Learning:
  - Everything becomes adaptive.
  - No distinction between feature extractor and classifier.
  - A complex and non-linear system trained from raw data.

- But…HOW???

Intuition behind Deep Neural Nets



About Deep Neural Nets

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Multi-task learning
- Transfer learning in deep learning

and a lot more (e.g., deep reinforcement learning, etc.)

That’s all for today.
Any Questions?

加簽結果公布