The National Higher School of Artificial Intelligence
المدرسة الوطنية العليا للذكاء الاصطناعي

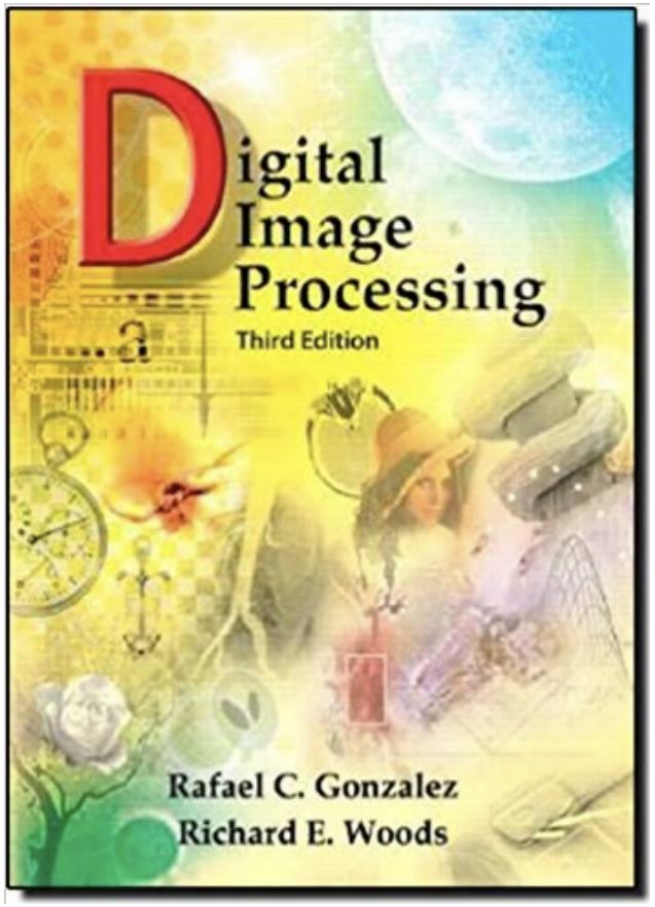
Course
« *Computer Vision*»

Sid-Ahmed Berrani

2024-2025

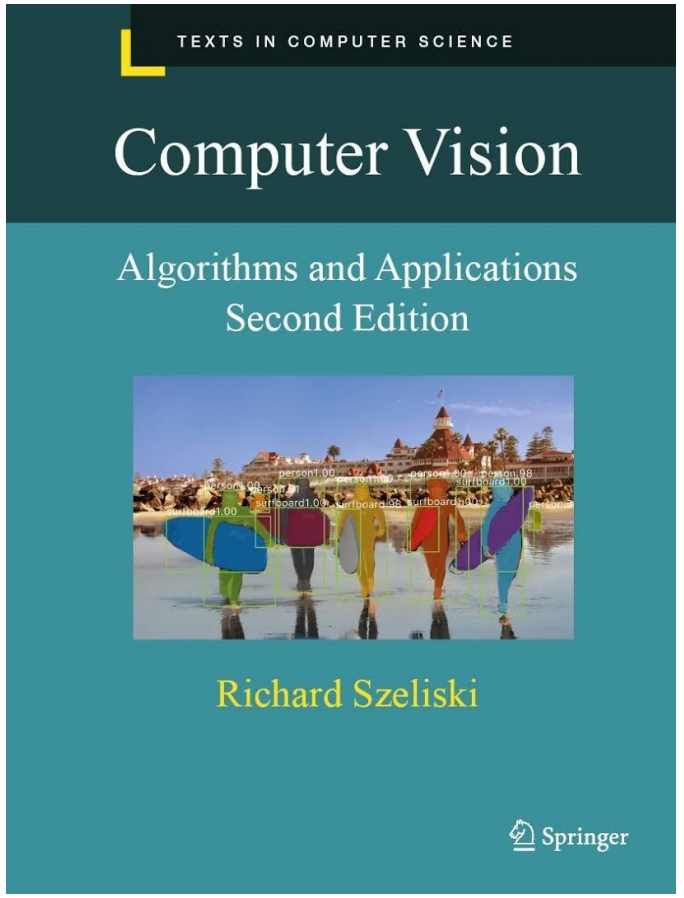
References

1. R. C. Gonzalez e R.E. Woods. Digital Image Processing (3rd edition). Pretience Hall.



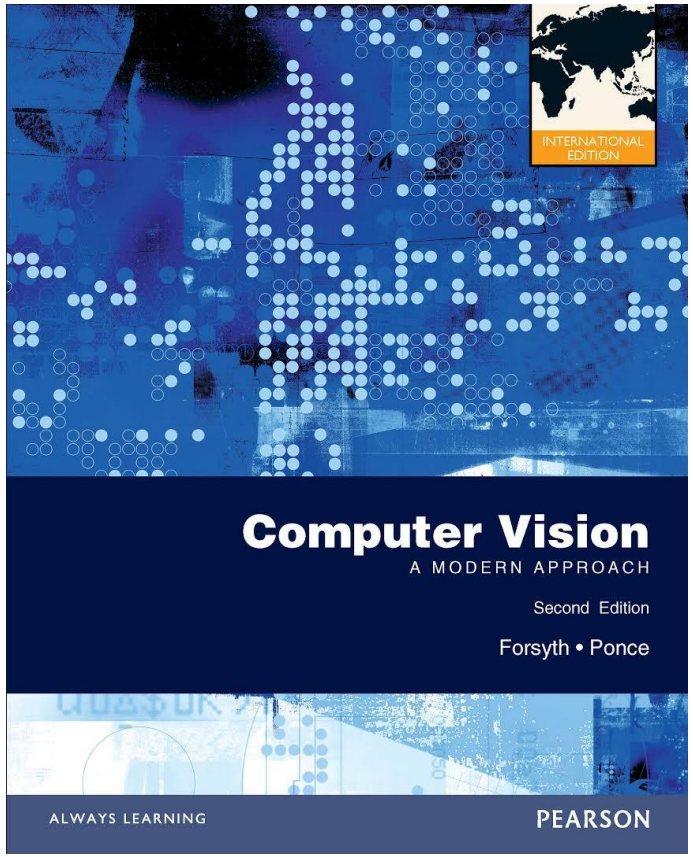
References

2. R. Szeliski. Computer Vision Algorithms and Applications (2nd edition). Springer.



References

3. D. Forsyth, J. Ponce. Computer Vision – A Modern Approach (2nd edition). Pearson.



Introduction

- Vision is the most powerful sense that allow us to interact with the physical world, without physical contact.
- Computer vision is the enterprise of building machines that can see.
- Computer vision is an enterprise that uses statistical methods to disentangle data using models constructed with the aid of geometry, physics, and learning theory.

Introduction

- Computer vision as a field is an *intellectual frontier*: Exciting and disorganized.
- Many useful ideas have no theoritical grounding
- Some theories are useless in practice.

⇒ *Difficult to give a concise definition since the area spans multiple different problems*

Introduction

- **How we will consider compter vision in this course:**

  "A set of computational techniques aiming at estimating or making explicit the geometric and dynamic properties of the 3D world from digital images."

⇒ *We will prioritize a practical approach while introducing the fundamental principles and concepts.*

Computer vision – different levels

1. **Low (Image processing)**
   - Image restoration
   - Contrast enhancement
   - Noise reduction

image → **image**

2. **Medium**
   - Image segmentation
   - Shape recognition

image → **Features/attributes/ descriptors**

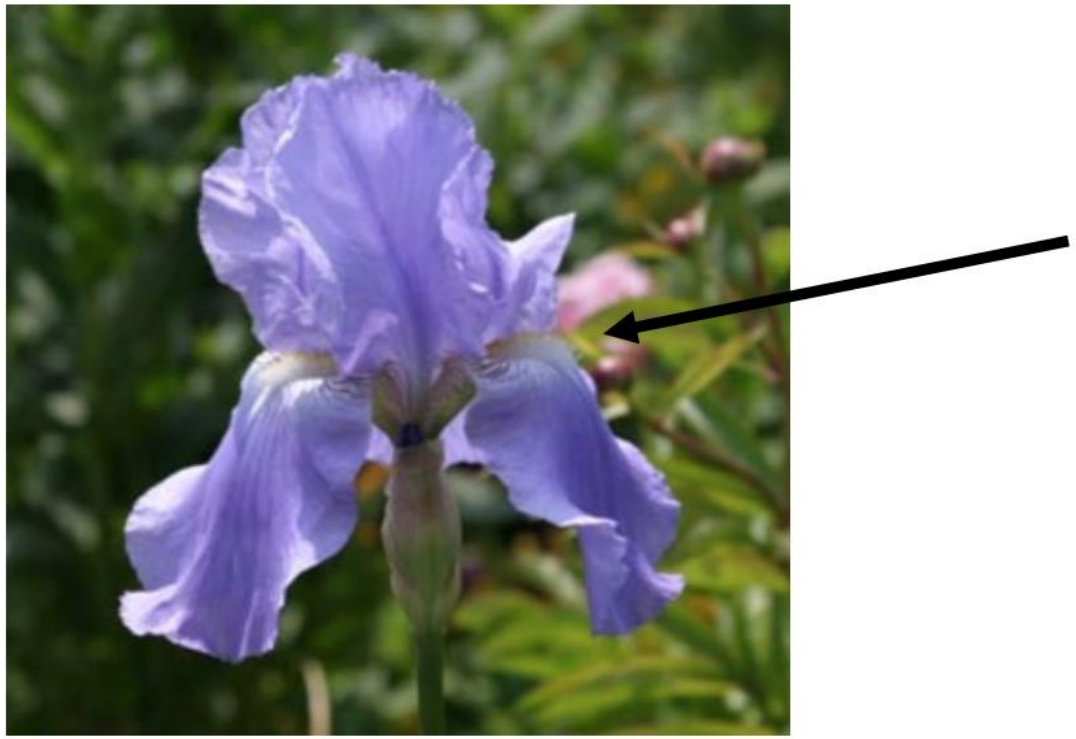
3. **High**
   - Scene understanding

image → **Concepts**

Introduction

- A human has no problem interpreting the subtle variations in translucency and shading in this photo and to perfectly segmenting the object from its background.

Introduction



Easy to distinguish the flower in the foreground from the leaves in the background

From a signal point of view, there are just variations in color and brightness

Introduction

- A human has no problem analyzing the photo, counting the number of people, classifying them (female/male), estimating the ages of people, recognizing people they know…

Introduction

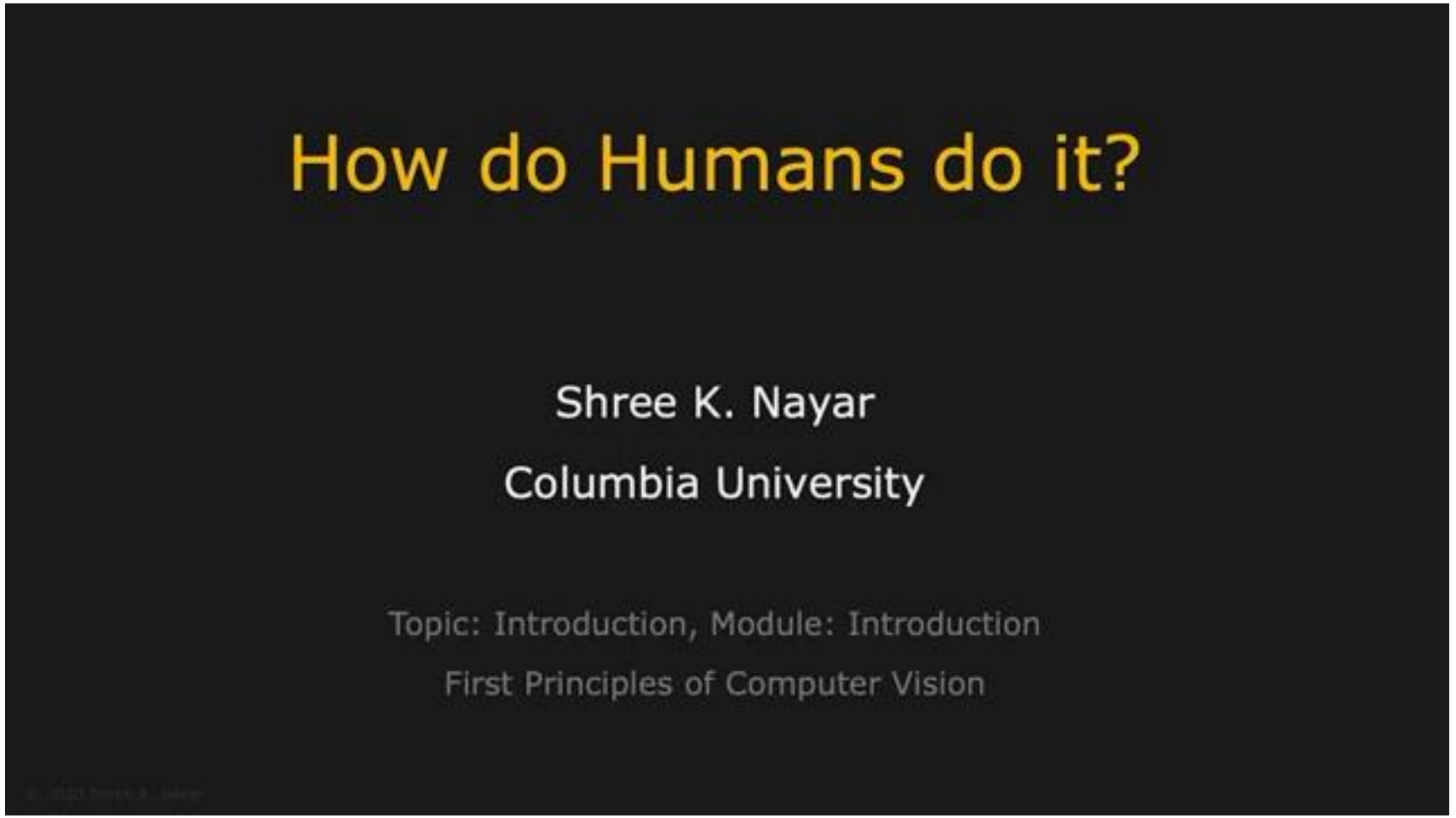
- But for machines, it is another story!
- These basic tasks are very challenging.

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Introduction: Human vision



Introduction

- Perceptual psychologists have spent decades trying to understand human visual system works.
- One idea is to make use of « optical illusions ».
- Researchers or scientists are able to create optical illusions to study and analyze some aspects of how the brain processes visual information.

Introduction – Optical illusions

- *« Optical illusions don't "trick the eye", nor "fool the brain", nor reveal that "our brain sucks", … but are fascinating! They also teach us about our visual perception and its limitations. »* - Michael Bach.

  https://michaelbach.de/ot/

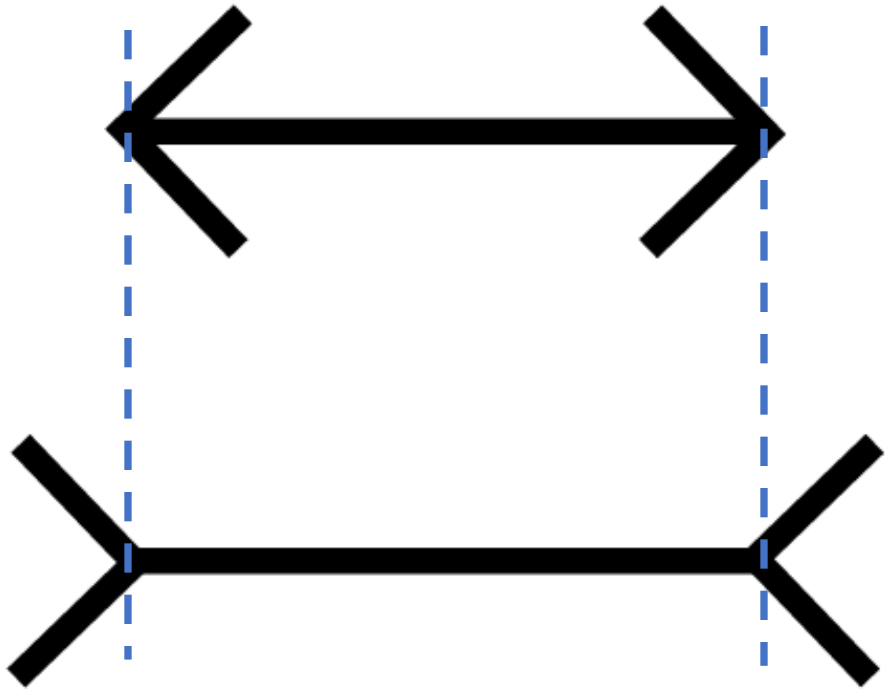
- **Example 1:** Motion Aftereffect (Waterfall Illusion)

  https://michaelbach.de/ot/mot-adapt/index.html

Introduction – Optical illusions

- **Example 2:** Müller-Lyer Illusion (1889)

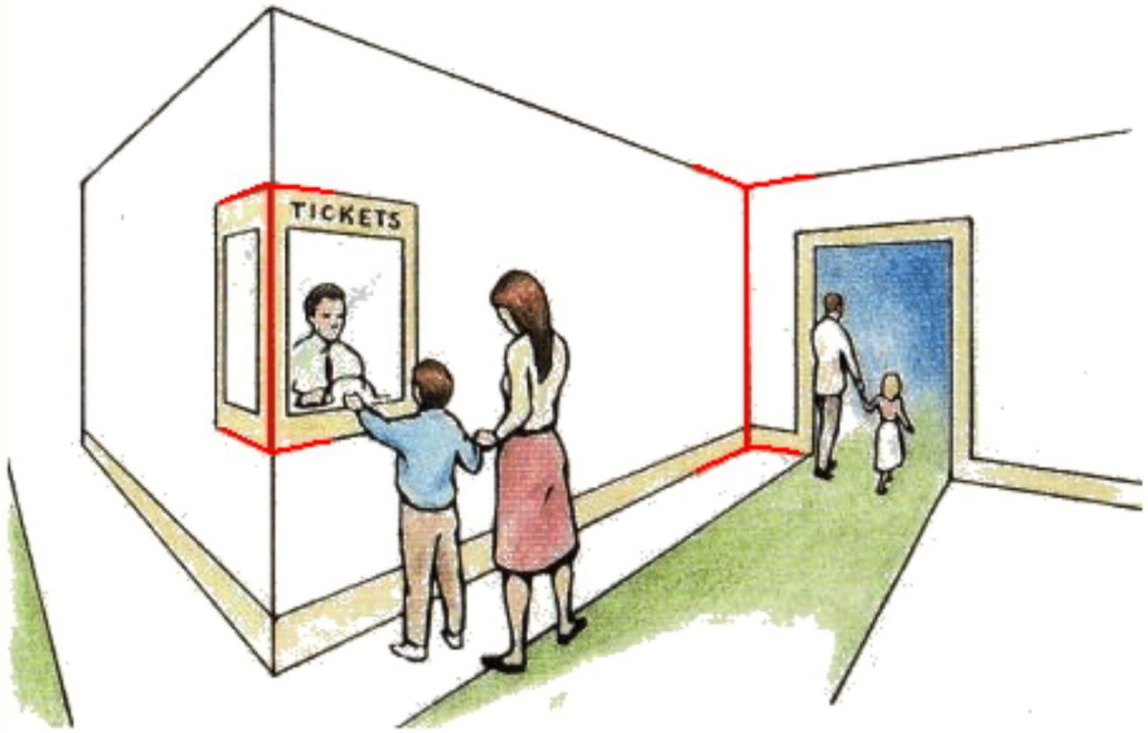
  https://michaelbach.de/ot/sze-muelue/index.html



Introduction – Optical illusions

- **Example 2:** Müller-Lyer Illusion (1889)

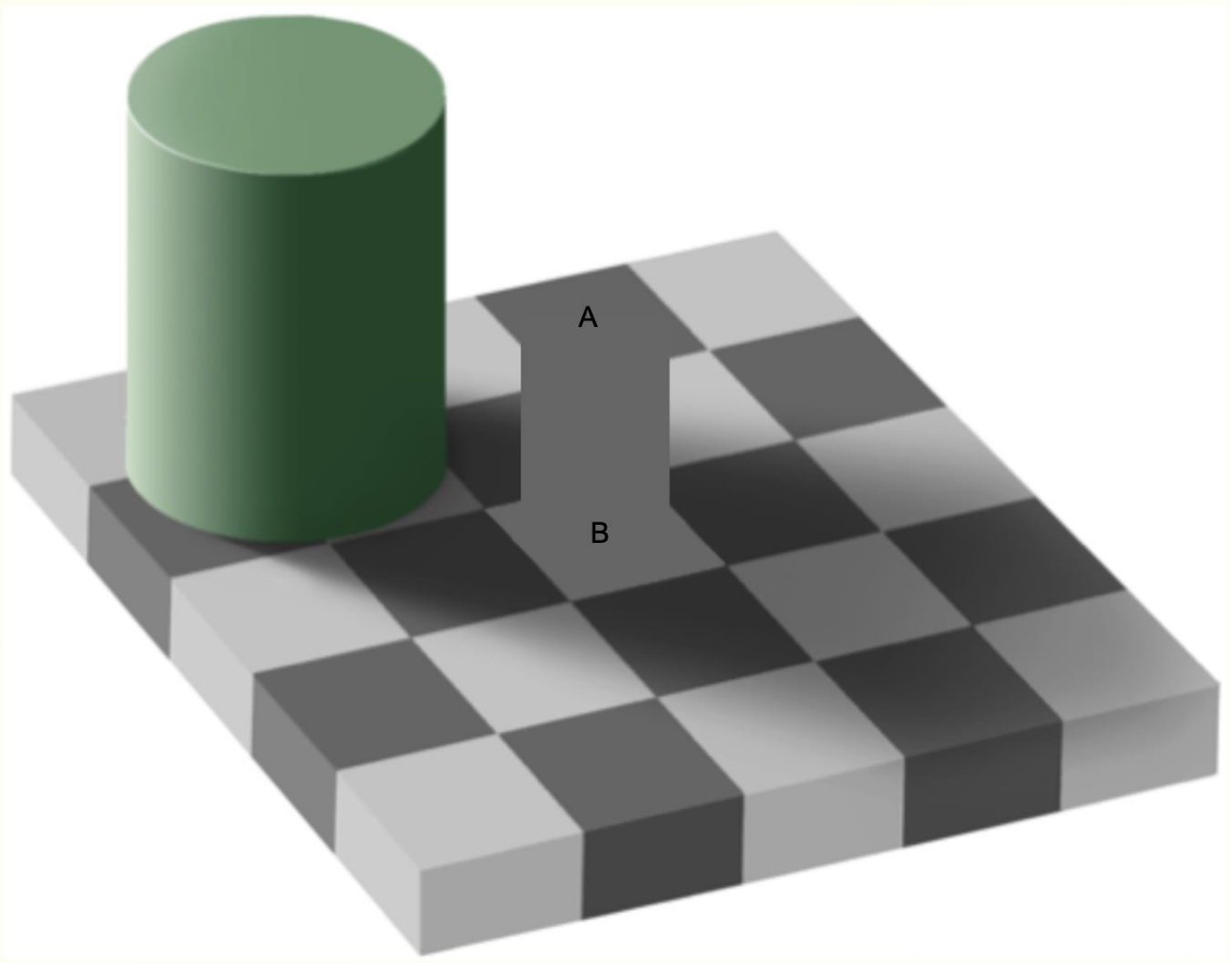
The brain assumes the "angles in" configuration to be closer, computes size constancy on it, and –given identical retinal size of the two angle arrangements – concludes that the "angle in"-line is shorter.



Introduction – Optical illusions

- **Example 3:** Adelson's Checker-shadow illusion

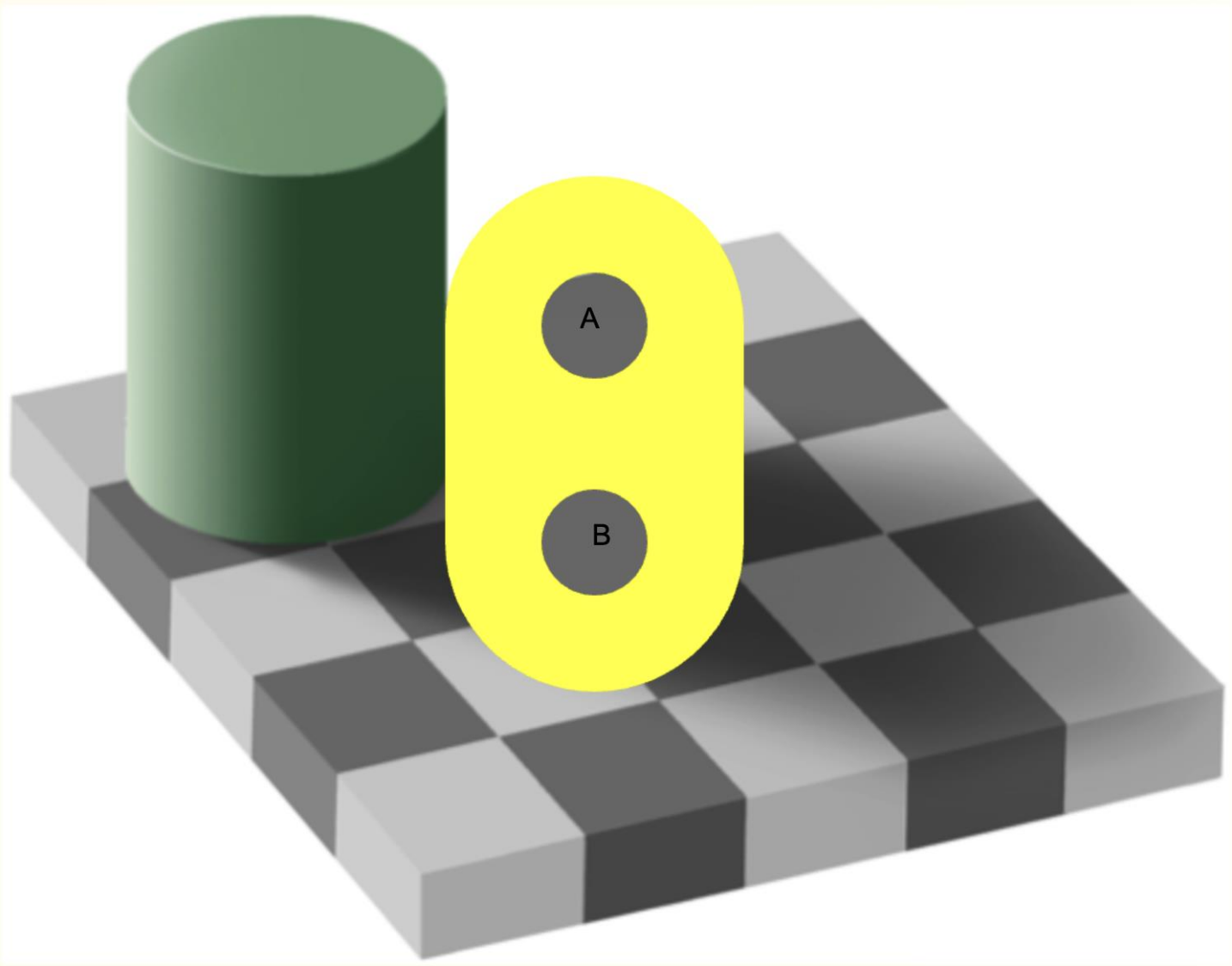
https://michaelbach.de/ot/lum-adelsonCheckShadow/index.html



Introduction – Optical illusions

- **Example 3:** Adelson's Checker-shadow illusion

https://michaelbach.de/ot/lum-adelsonCheckShadow/index.html



Introduction – Optical illusions

- **Example 3:** Adelson's Checker-shadow illusion

https://michaelbach.de/ot/lum-adelsonCheckShadow/index.html

- Explanation

The visual system needs to determine the color of objects in the world. In this case the problem is to determine the gray shade of the checks on the floor. Just measuring the light coming from a surface (the luminance) is not enough: a cast shadow will dim a surface, so that a white surface in shadow may be reflecting less light than a black surface in full light. The visual system uses several tricks to determine where the shadows are and how to compensate for them…

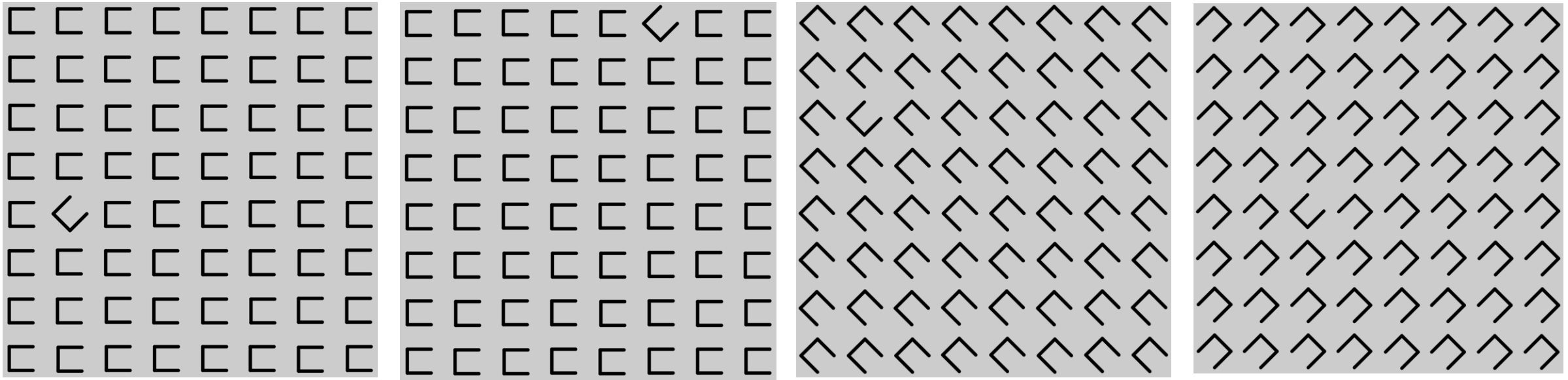
https://persci.mit.edu/gallery/checkershadow/description

Introduction – Optical illusions

- **Example 4:** the pop-out effect

https://michaelbach.de/ot/cog-popout/index.html

Search for the one U that has an orientation than all the others.



Introduction – Optical illusions

- **Example 4:** the pop-out effect

The **pop-out effect** in human visual perception refers to how certain objects or features in a visual scene immediately stand out and grab attention because they are different from their surroundings. This happens effortlessly and without the need to actively search for them.

The pop-out effect can be explained by the way the brain processes visual information, particularly through mechanisms in the visual cortex and attentional systems.

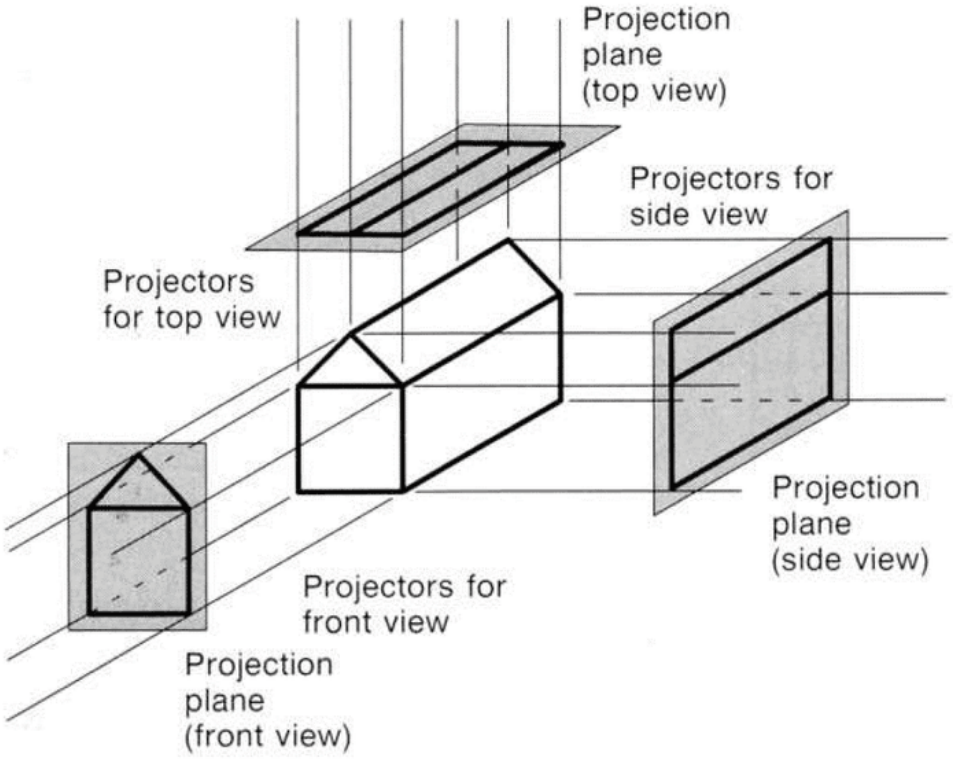
The pop-out effect occurs because the brain is equipped with specialized neurons, parallel processing, and attentional mechanisms that prioritize unique or salient visual features. This automatic, efficient system helps us navigate and respond to our surroundings with minimal effort.

Introduction – Definitions

- A human perceives the 3D structure of the world with apparent ease.
- Researchers study how the human visual system works.
- Researchers have been developing numerous techniques for decades to emulate human vision.
- Why is vision so difficult?
  - Mainly because is an inverse problem.
  - We seek to recover some unknowns given insufficient information to fully specify the solution.

Introduction – Definitions

- Computer vision is difficult because it often involves solving an inverse problem:
  - The goal is to infer the real-world properties of a scene (e.g., objects, depth, lighting) from a 2D image or set of images.
  - The challenge arises because a 2D image contains limited and ambiguous information, making it impossible to uniquely determine the exact 3D world that created it without additional assumptions.



CV vs. Computer Graphics

Computer graphics

3d to 2d implies information loss

Computer Vision

Need for models

Raster image

- A house
- Some polygons
- Lines, edges

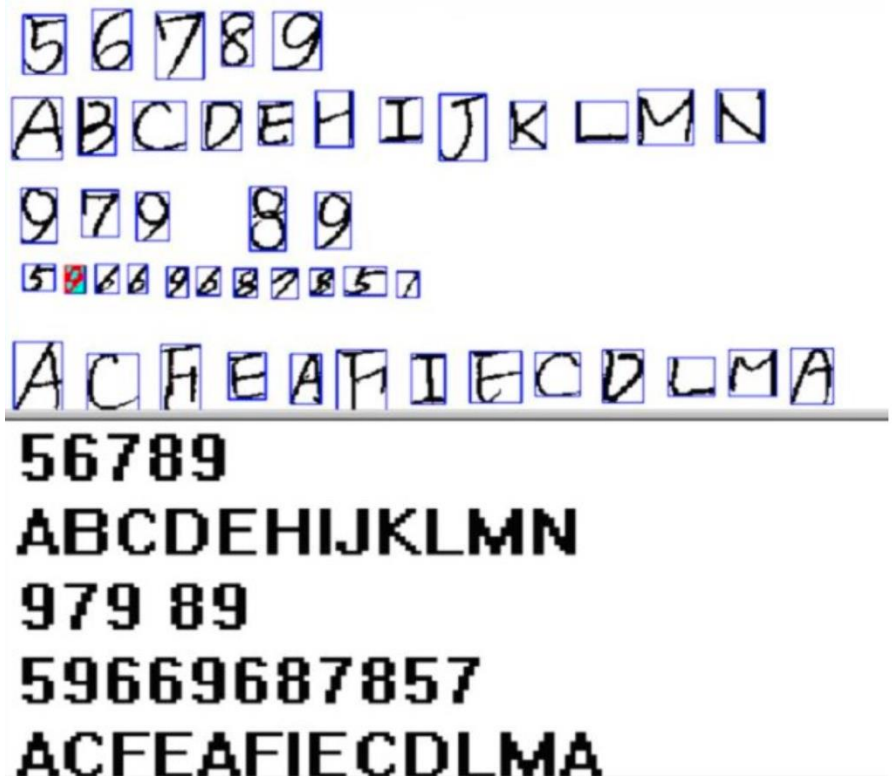
CV – Applications

- The good news, CV is being used today in a wide variety of real-world applications:
  - OCR
  - Machine inspection
  - Medical imaging
  - Self-driving cars
  - Surveillance
  - Biometrics (face, iris, fingerprints…)
  - Image and video fingerprinting
  - Content-based image/video retrieval
  - Similarity search
  - …

+ Consumer-level applications.

CV – Applications

- Optical Character Recognition (OCR)



CV – Applications

- Video OCR

Overview of the ALIF Dataset

هل دخلت تونس

في مسلسل الاغتيالات السياسية ؟

S. Yousfi, S.-A. Berrani, C. Garcia. ***Arabic text detection in videos using neural and boosting-based approaches: Application to video indexing****. IEEE ICIP 2014.*

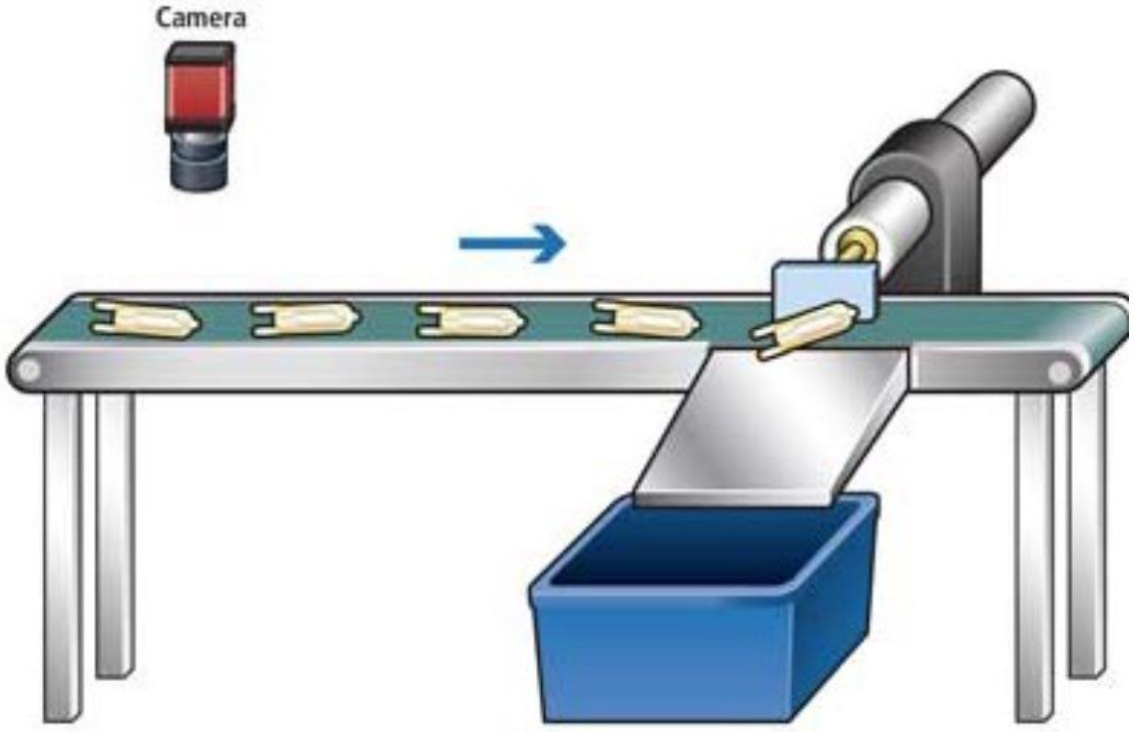
S. Yousfi, S.-A. Berrani, C. Garcia. ***Boosting-based approaches for Arabic text detection in news videos****. DAS 2014.*

S. Yousfi, S.-A. Berrani, C. Garcia. ***ALIF: A Dataset for Arabic Embedded Text Recognition in TV Broadcast.*** *ICDAR 2015. S. Yousfi, S.-A. Berrani, C. Garcia.* ***Deep Learning and Recurrent Connectionist-based Approaches for Arabic Text Recognition in Videos****. ICDAR 2015.*

S. Yousfi, S.-A. Berrani, C. Garcia. ***Contribution of recurrent connectionist Language Models in Improving LSTM-Based Arabic Text Recognition in Videos****. Pattern Recognition 64 : 245-254 (2017).*

CV – Applications

- Machine inspection

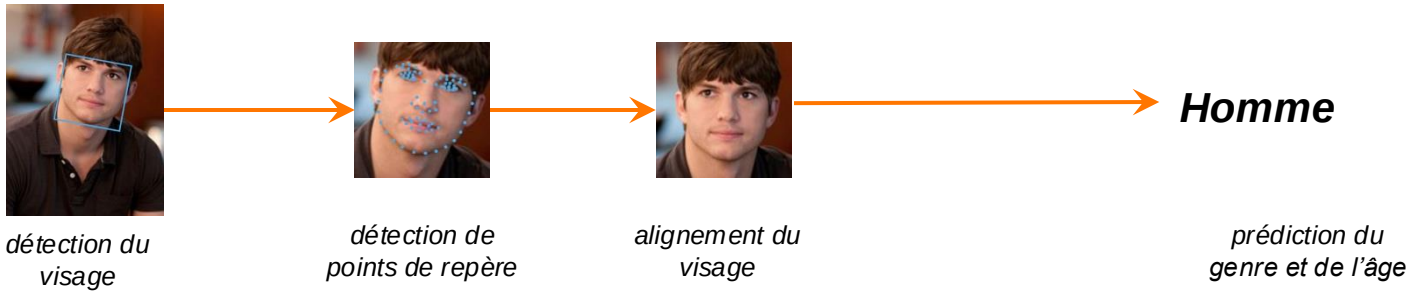


CV – Applications

- Surveillance and tracking

CV – Applications

- Soft biometrics: Gender detection from face images



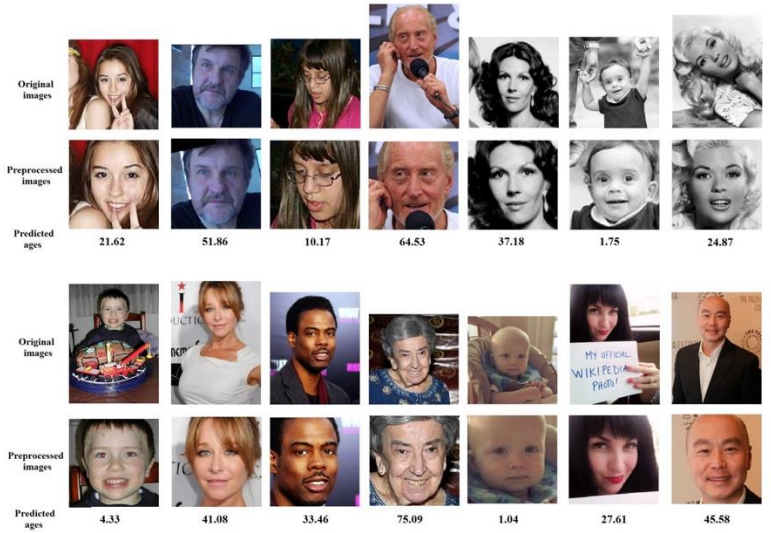
- « Learned vs. Hand-Crafted Features for Pedestrian Gender Recognition » par G. Antipov, S.-A. Berrani, N. Ruchaud et J.-L. Dugelay. **ACM Conference on Multimedia 2015**.
- « Minimalistic CNN-Based Ensemble Model for Gender Prediction from Face Images » par G. Antipov, S.-A. Berrani et J.-L. Dugelay. **Pattern Recognition Letters 2015**.
- « The impact of privacy protection filters on gender recognition » par N. Ruchaud, G. Antipov, P. Korshunov, J.-L. Dugelay, T. Ebrahimi et S.-A. Berrani. **SPIE Optical Engineering 2015**.

CV – Applications

- Soft biometrics: Apparent age estimation from face images

Position	Team	ϵ-score[a]
1	**OrangeLabs (our team)**	**0.2411**
2	palm_seu	0.3214
3	cmp+ETH	0.3361
4	WYU_CVL	0.3405
5	ITU_SiMiT	0.3668
6	Bogazici	0.3740
7	MIPAL_SNU	0.4569
8	DeepAge	0.4573

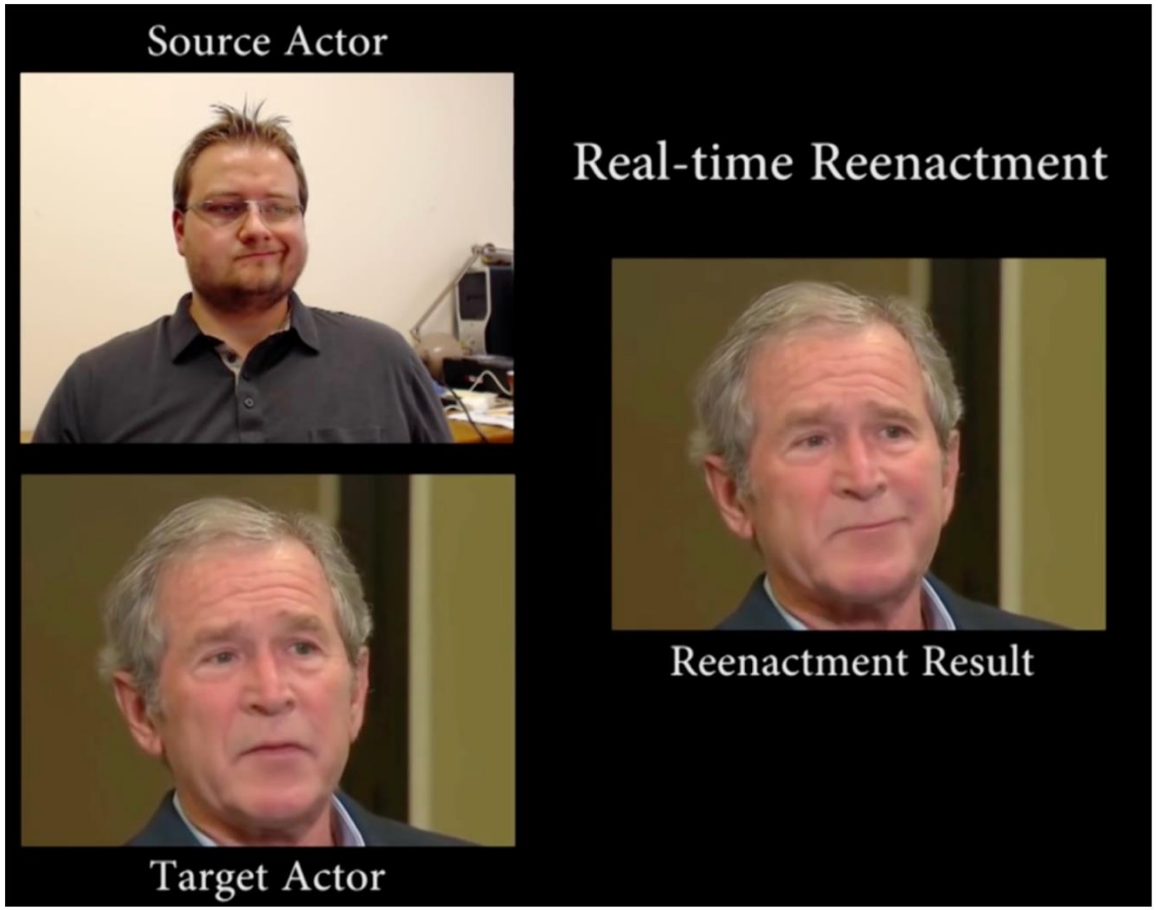
ChaLearn LAP and FotW Challenge and Workshop @ CVPR2016
Winner of Track 1: Apparent Age Estimation



- « Effective Training of Convolutional Neural Networks for Face-Based Gender and Age Prediction » par G. Antipov, M. Baccouche S.-A. Berrani, J.-L. Dugelay. **Pattern Recognition, 2017**.
- « Apparent Age Estimation from Face Images Combining General and Children-Specialized Deep Learning Models » par G. Antipov, M. Baccouche, S.-A. Berrani, J.-L. Dugelay. **Computer Vision and Pattern Recognition Workshop 2016**.

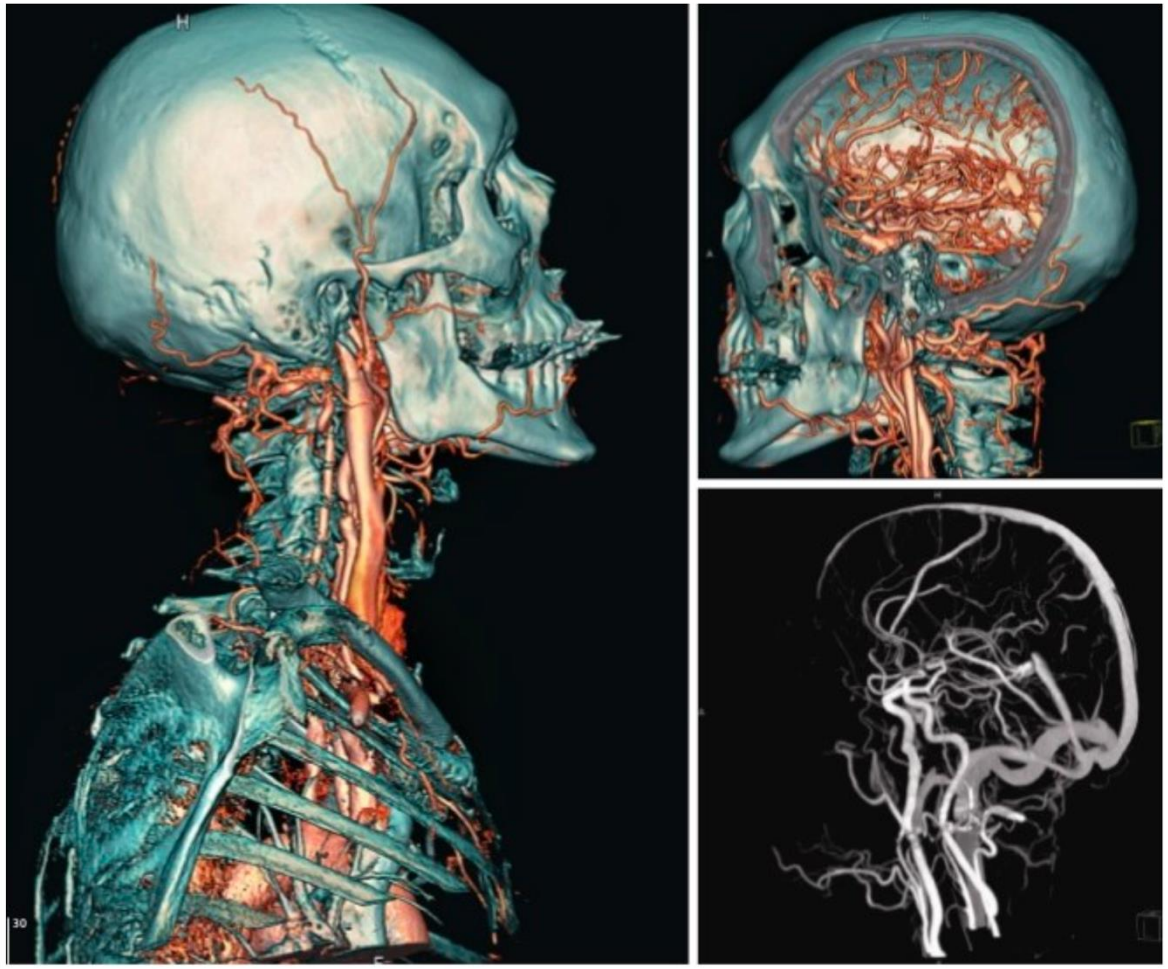
CV – Applications

- Face capture and reenactment



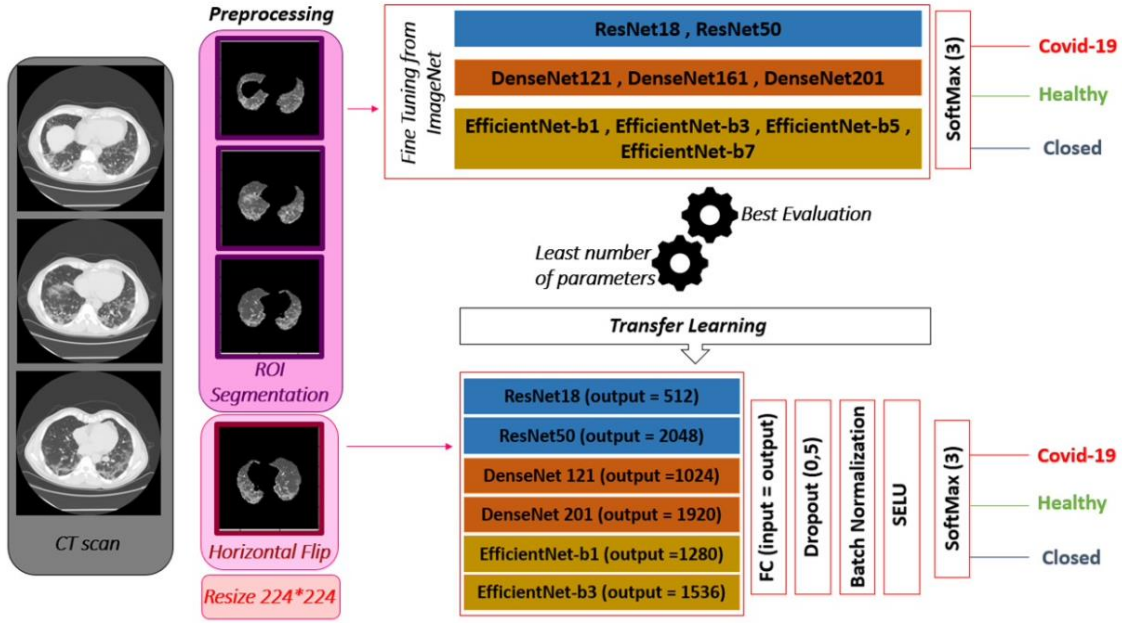
CV – Applications

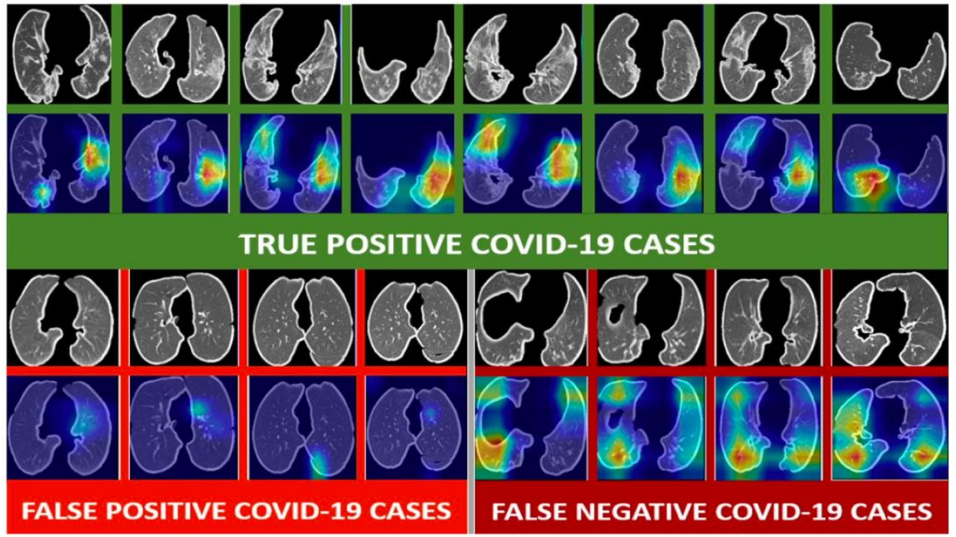
- Medical image analysis



CV – Applications

- Medical image analysis: Covid-19 diagnosis from CT images

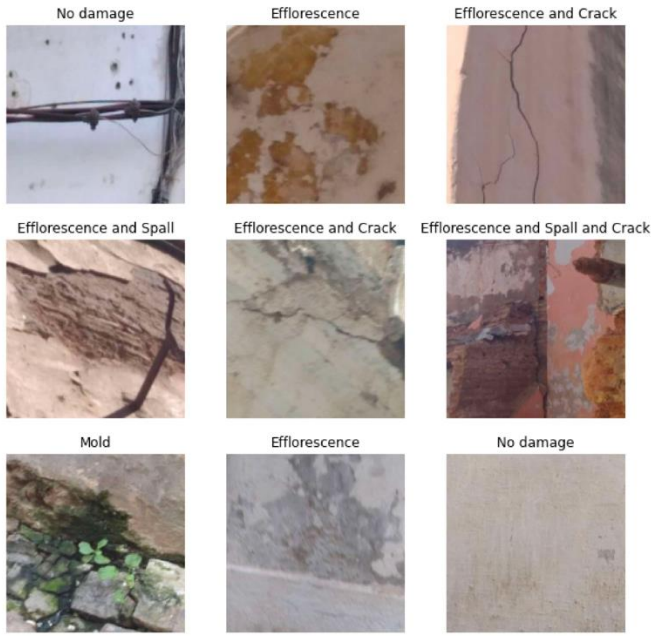




- I. Hadj Bouzid, S. Yahiaoui, A. Lounis, S.-A. Berrani, H. Belbachir, Q. Naïli, M. E. Abdi, K. Bensalah, D. Belazzougui. **DIAG a diagnostic web application based on lung CT Scan images and deep learning.** *The 31st Medical Informatics Europe Conference.* May 2021.
- I. Hadj Bouzid, S.-A. Berrani, S. Yahiaoui, B. Ahror, D. Belazzougui, M. Djouad, K. Bensalah, H. Belbachir, Q. Naïli, M. E. Abdi. **Deep Learning-based Covid-19 Diagnosis: A Thorough Assessment with a Focus on Generalization Capabilities.** *Eurasip Journal on IVP 2024.*

CV – Applications

- Surface Damage Identification for Heritage Site Protection: A solution for the Kasbah of Algiers



A world heritage city
+ six centuries old

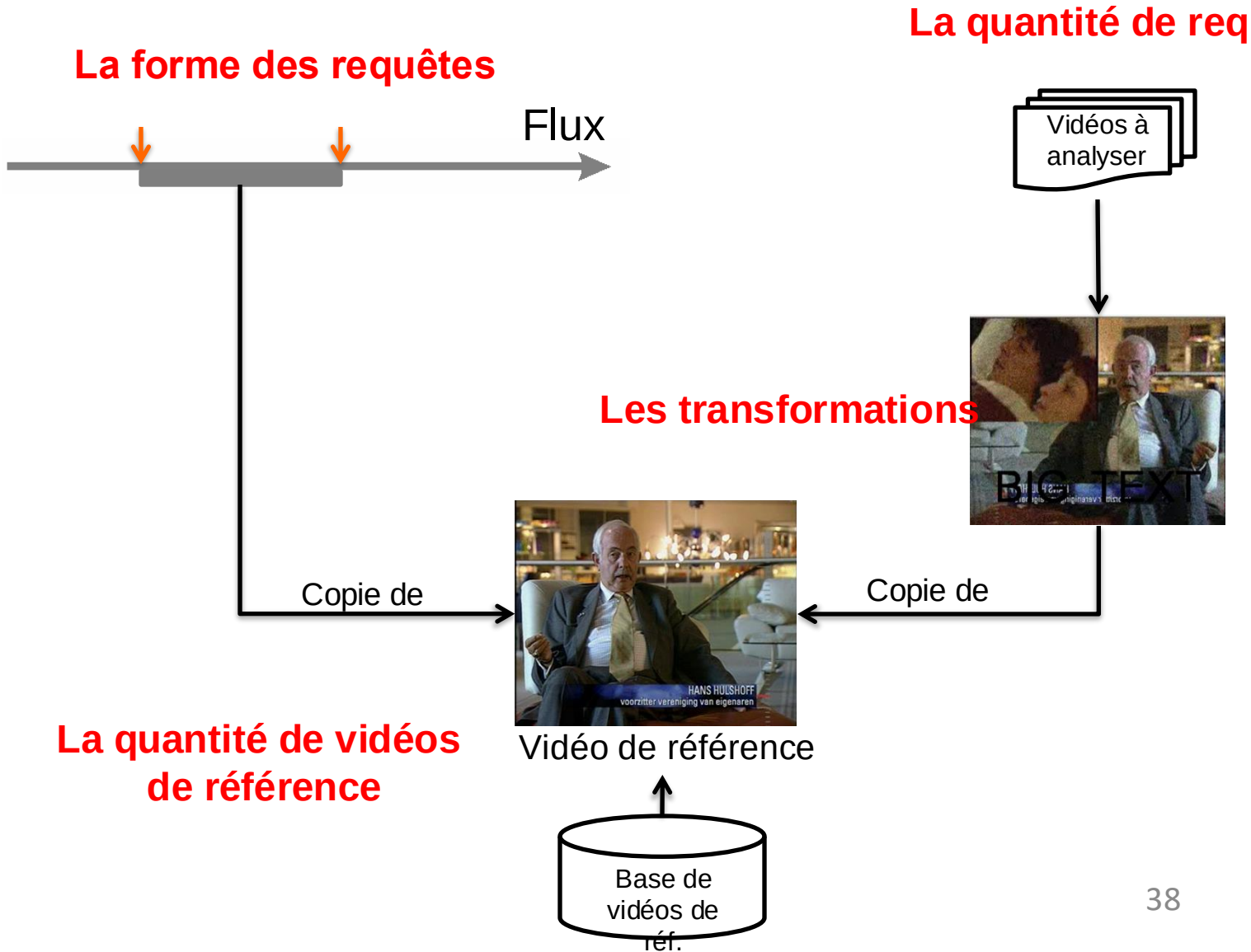
S. Meklati, K. Boussora, M. E. Abdi, S.-A. Berrani. **Surface Damage Identification for Heritage Site Protection: A Mobile Crowd-sensing Solution Based on Deep Learning.** *ACM Journal on Computing and Cultural Heritage 16(2): 1-24 (2023).*

CV – Applications

- Self-driving cars

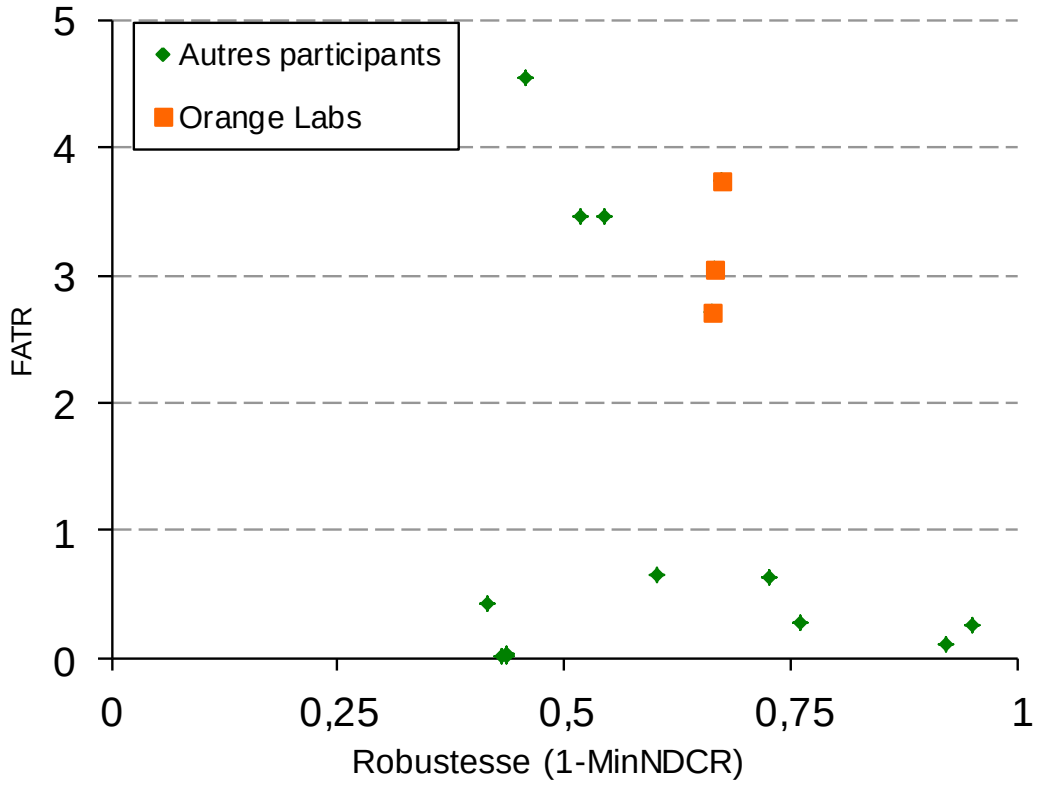
CV – Applications

- Image/video copy detection: Fingerprinting



CV – Applications

- Image/video copy detection: Fingerprinting



Publications:

N. Gengembre, S.-A. Berrani. The Orange Labs Real Time Video Copy Detection System -- TrecVid 2008 Results. TrecVID Workshop 2008.

N. Gengembre, S.-A. Berrani. A Probabilistic Framework for Fusing Frame-based Searches within a Video Copy Detection System. ACM CIVR 2008.

N. Gengembre, S.-A. Berrani, P. Lechat. Adaptive Similarity Search in Large Databases -- Application to Video Copy Detection.CBMI 2008.

Patents:

1. EP2677764
2. WO20100116093
3. WO2010045699
4. WO2009095616
5. WO2009081016

CV – Applications

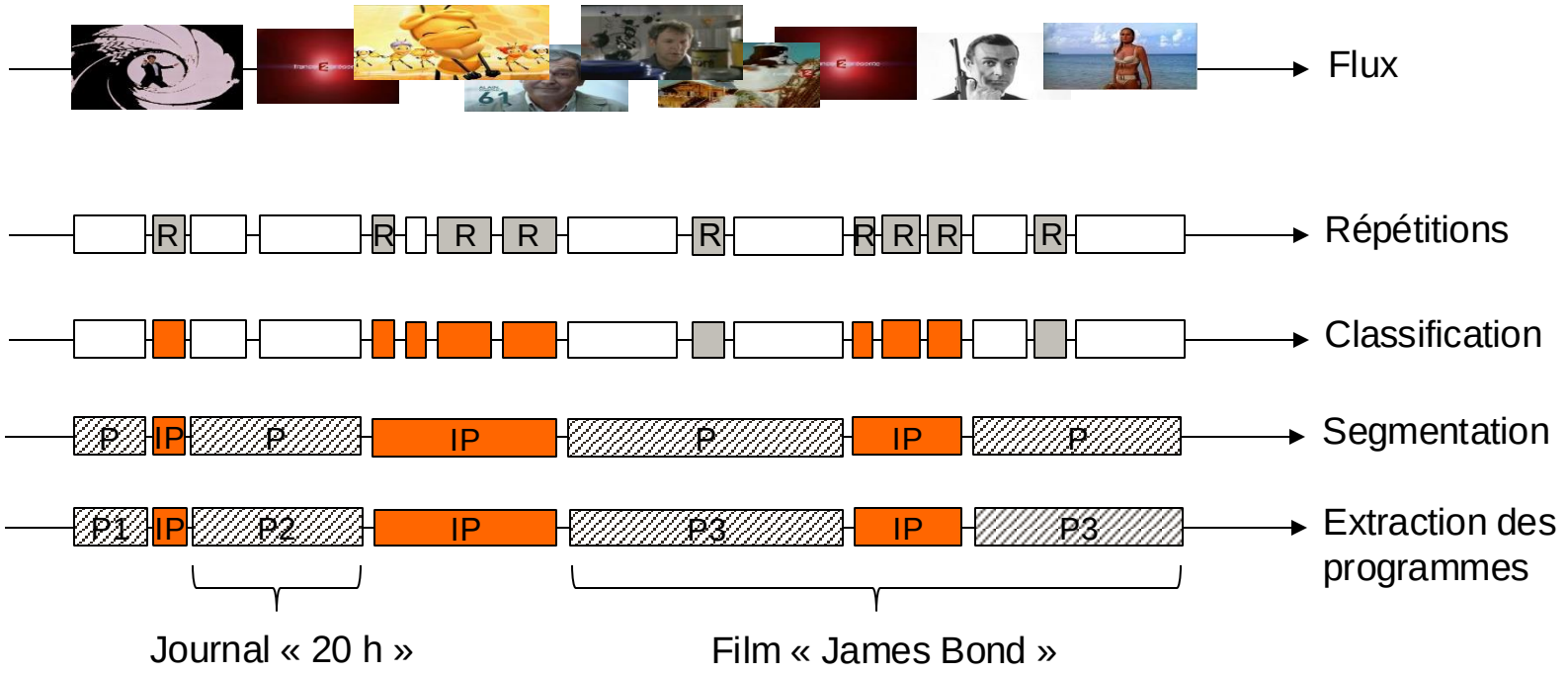
- Image/video copy detection: Fingerprinting

Copies

Originales

CV – Applications

- Video stream macro-segmentation and structuring



S.-A. Berrani, G. Manson, P. Lechat. A Non-Supervised Approach for Repeated Sequence Detection in TV Broadcast Streams. Signal Processing: Image Communication, 23(7):525-537, 2008.

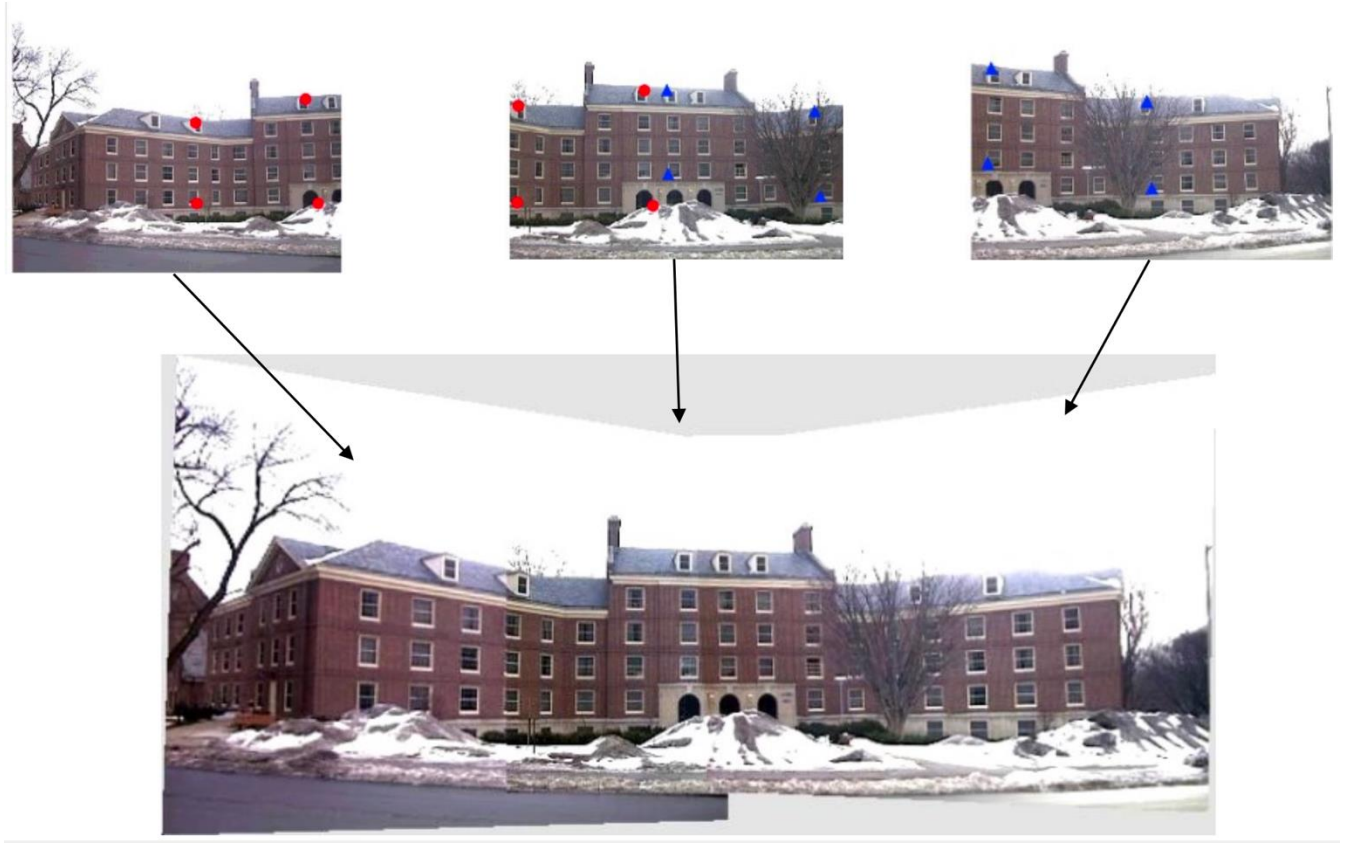
Patent : WO2008081143

Introduction – Applications

- Consumer-level applications:
  - Stitching
  - Image enhancement
  - Morphing
  - Face-based authentication
  - …

CV – Applications

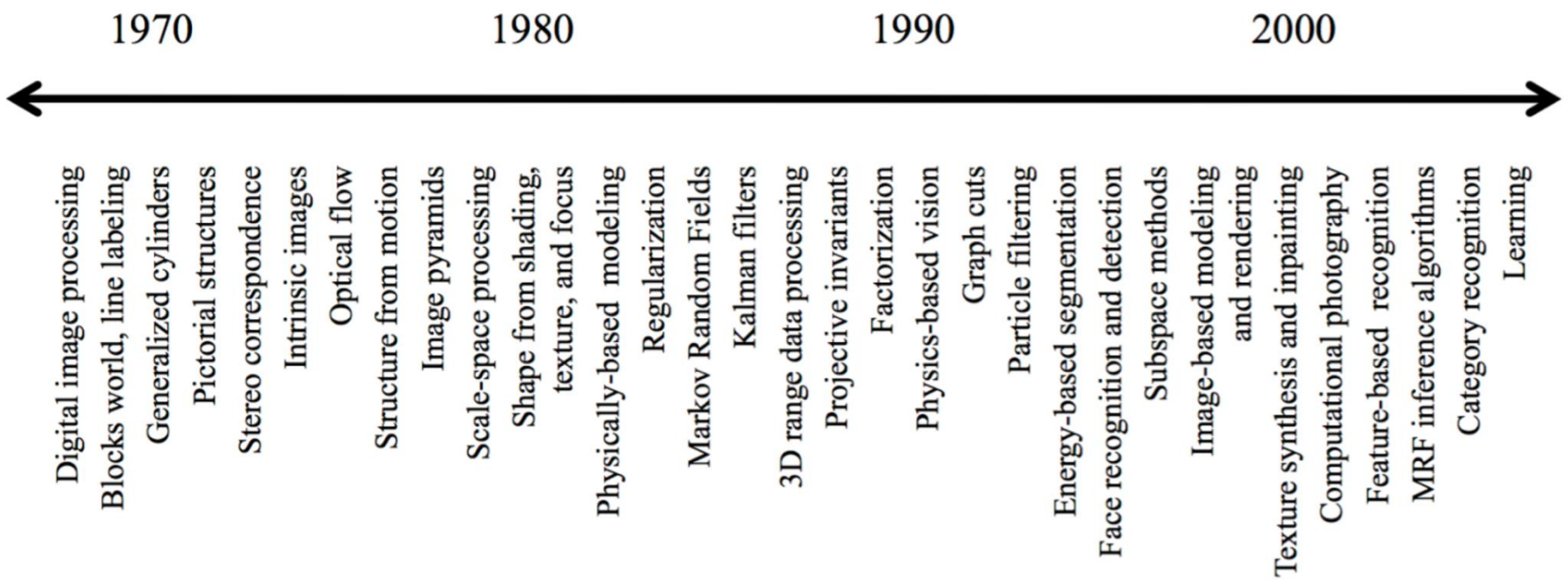
- Image Stitching



Computer Vision: A Brief History

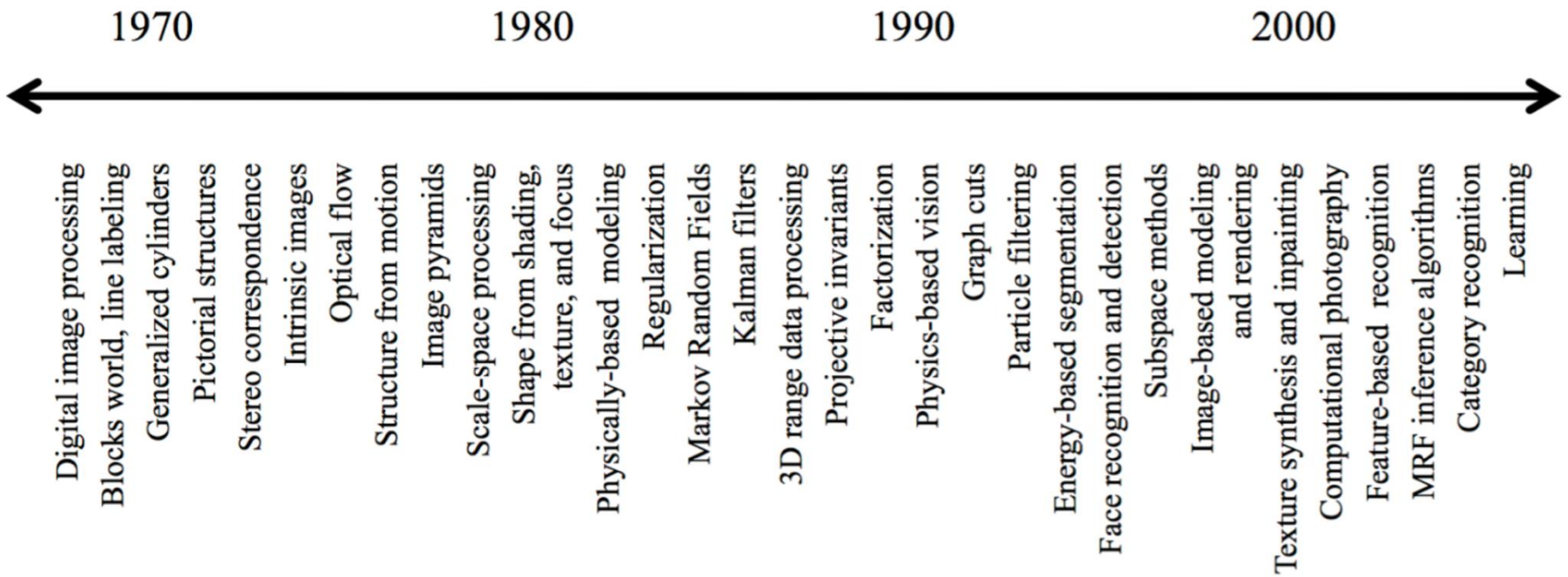
- **1970:** What started distinguishing computer vision from image processing were early attempts to infer 3D structure from images:

Lines extraction and labelling, stereo correspondences, optical flow, structure from motion



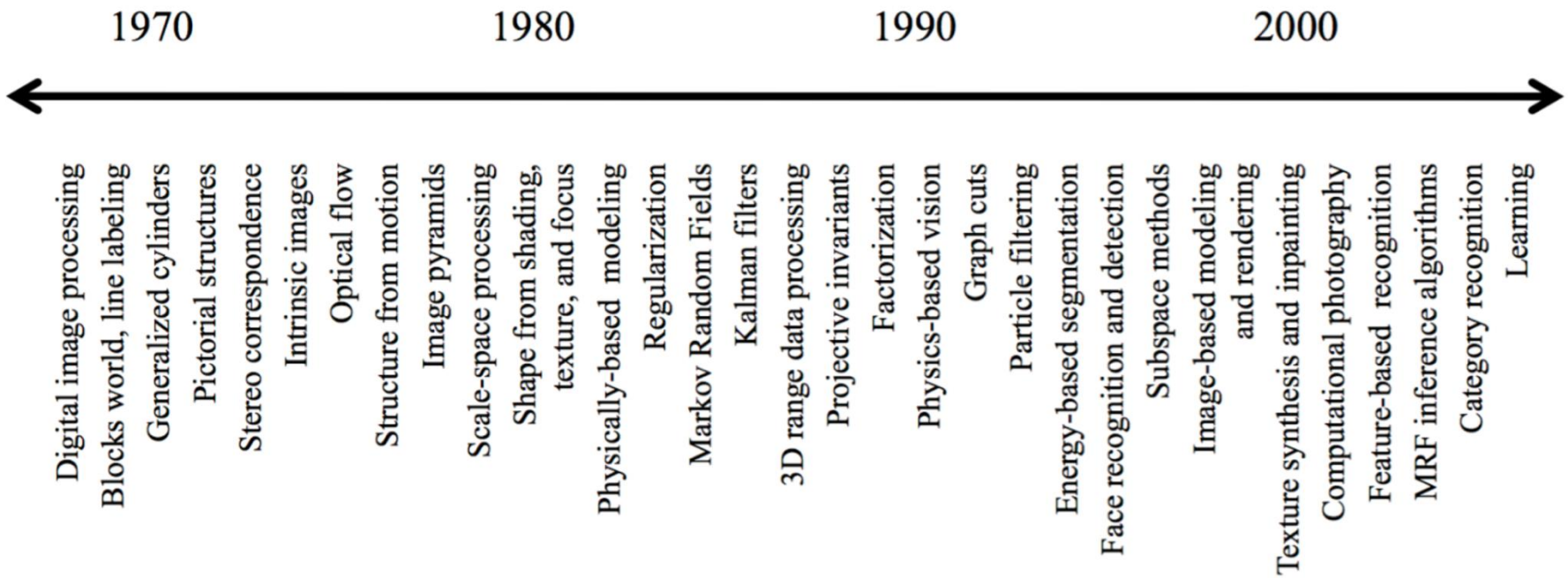
Computer Vision: A Brief History

- **1980:** attention was focused on more sophisticated mathematical techniques for performing quantitative image and scene analysis: Variational optimization, MRFs, Image Pyramids, 3D scanning…



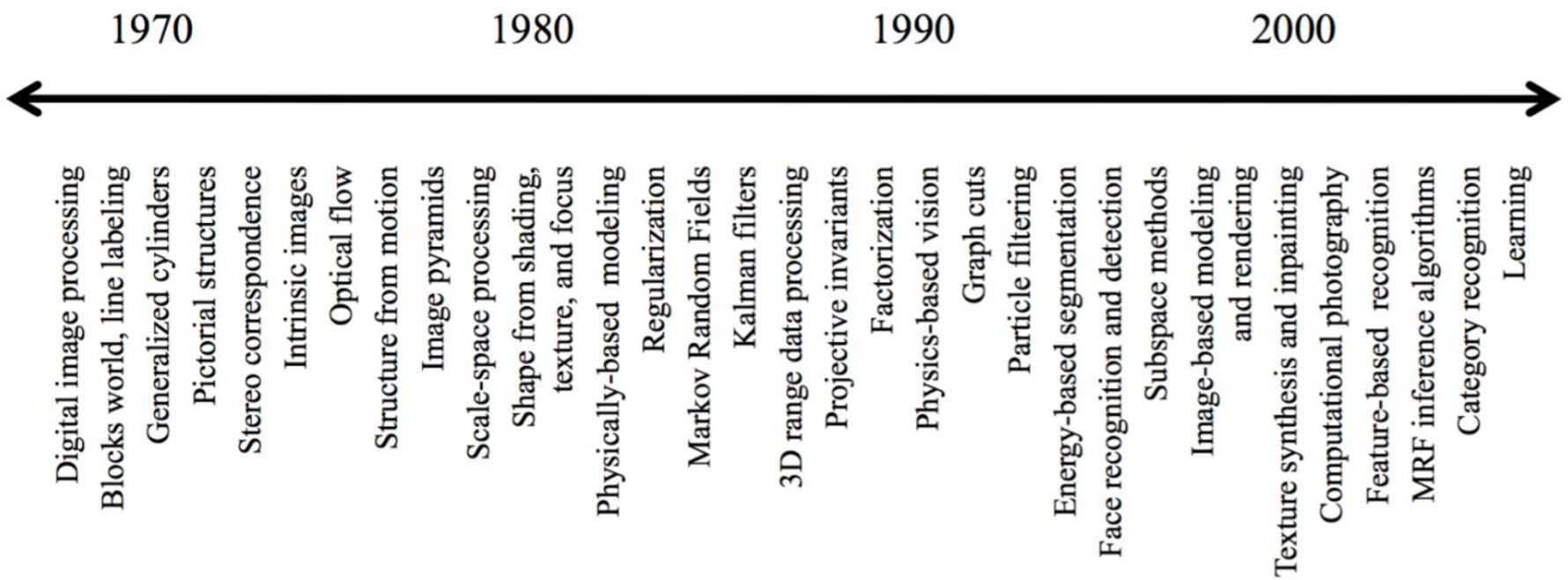
Computer Vision: A Brief History

- **1990:** While a lot of the previously mentioned topics continued to be explored, a few of them became significantly more active: projective invariants, multiview stereo, image segmentation…



Computer Vision: A Brief History

- **2000 - today:** Continuous advances of all the previous topics: SIFT features, texture synthesis, computational photography, learning…



Computer Vision: A Brief History

- **2012 - Today:** Deep learning