Vision Intelligence I: Introduction

Wei Qi Yan

Auckland University of Technology

Table of Contents

Vision Intelligence I: Introduction

Artificial Intelligence

- Perception (Observation)
- Learning
- Presentation
- Reasoning (Inference)

Relevant Subjects

- Computer Graphics
- Image and Video Processing
- Computer Vision
- Multimedia Computing
- Computational Linguistics
- Speech Recognition
- Natural Language Processing

Vision Intelligence I: Introduction

Computer Vision

- Motion analysis
- Stereo and monocular vision
- Contour, blob, silhouette
- Object segmentation
- Object detection
- Object recognition
- Object tracking
- Behavior analysis
- Visual event computing
- Intelligent navigation
- .....

Questions?

Questions?

Artificial Intelligence (AI) is related to:

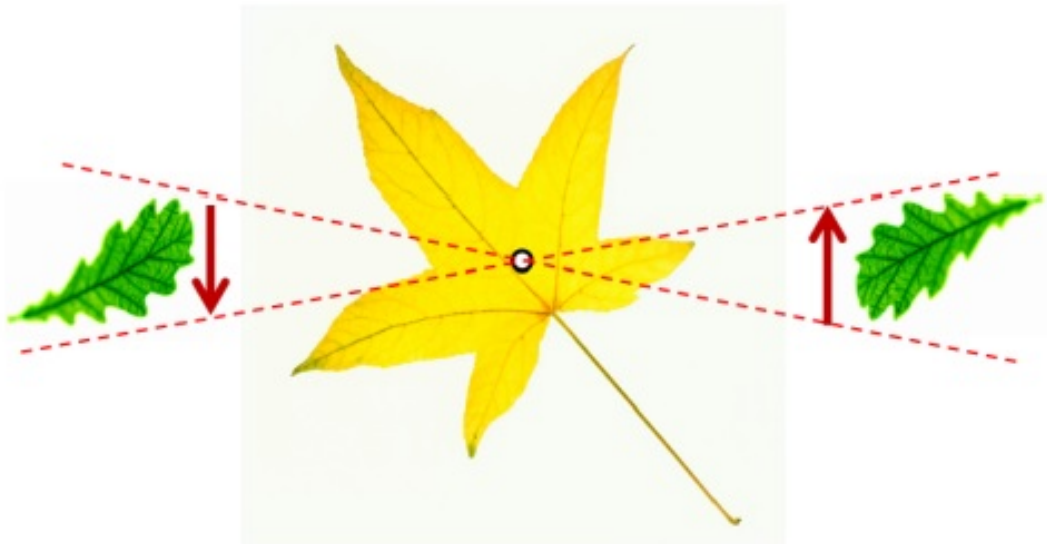
1. Perception, Learning, Presentation, Reasoning
2. Observation, Learning, Presentation, Reasoning
3. Perception, Learning, Presentation, Inference
4. None of the given options

The wrong answer is: ___

Questions?

Image Formation

- Geometry of image formation: Virtual
- Photometry of image formation: Real

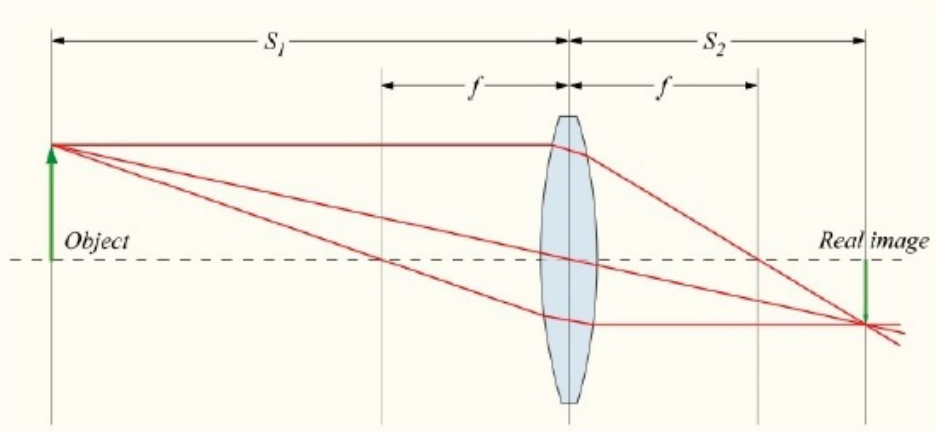


Lens System

A lens is much wider than a pinhole, enabling to let in more light. This is paid for by the fact that not all the scene can be in sharp focus at the same time.

$$\frac{1}{S_1} + \frac{1}{S_2} = \frac{1}{f}$$

where f is the focal length of the lens, S_1 is the object distance, S_2 is the image distance.



Color: Image Formation

- **Binary image**: $I(x, y) = (R, G, B)$, $R = G = B$, $R, G, B \in \{0, 2^b - 1\}$, b (e.g., $b = 8$) is the bit plane; $x = 1, 2, \cdots, width$; $y = 1, 2, \cdots, height$
- **Grayscale image**: $I(x, y) = (R, G, B)$, $R = G = B$, $R, G, B \in \{0, 1, 2, 3, \cdots, 2^b - 1\}$
- **Color image**: $I(x, y) = (R, G, B)$, $R, G, B \in \{0, 1, 2, 3, \cdots, 2^b - 1\}$
- **Real color image**: $I(x, y) = (R, G, B, \alpha)$, $R, G, B \in \{0, 1, 2, 3, \cdots, 2^b - 1\}$, $\alpha \in [0, 1]$

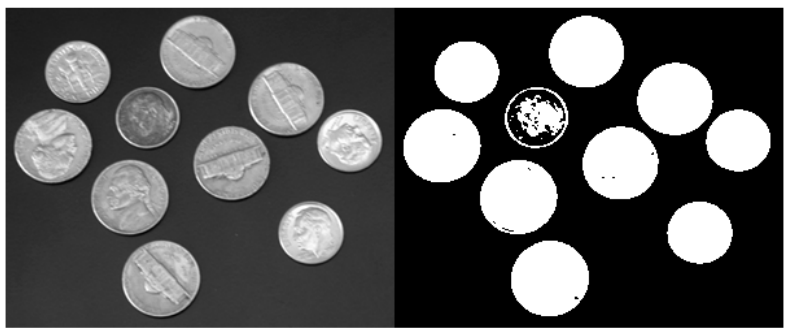
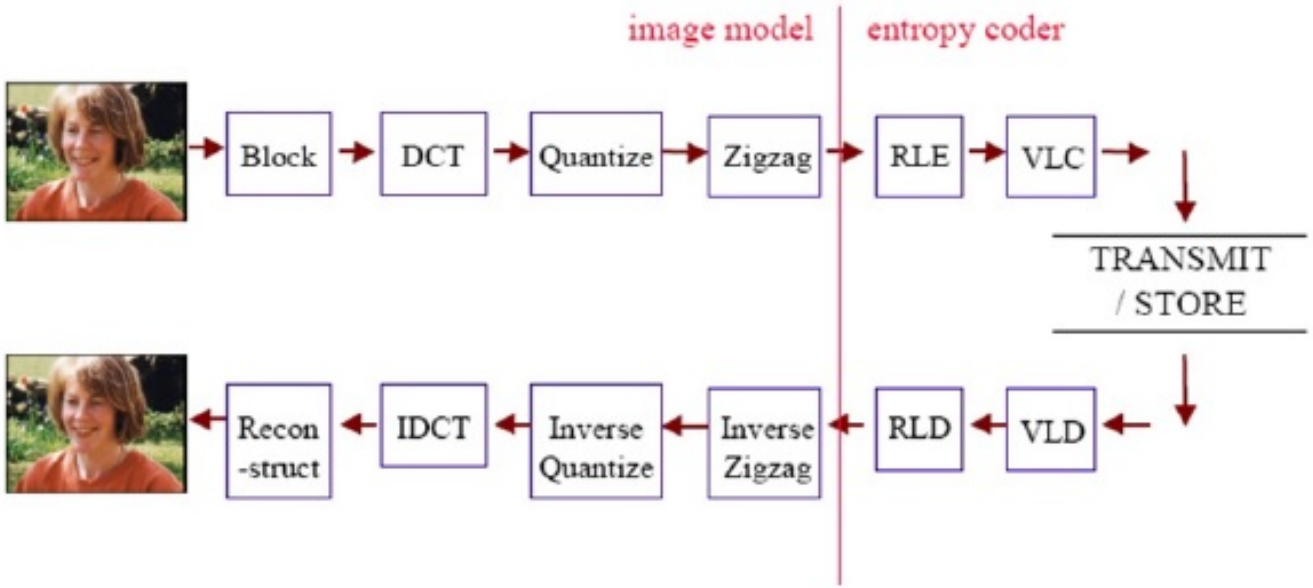


Image Set

JPEG Image Compression

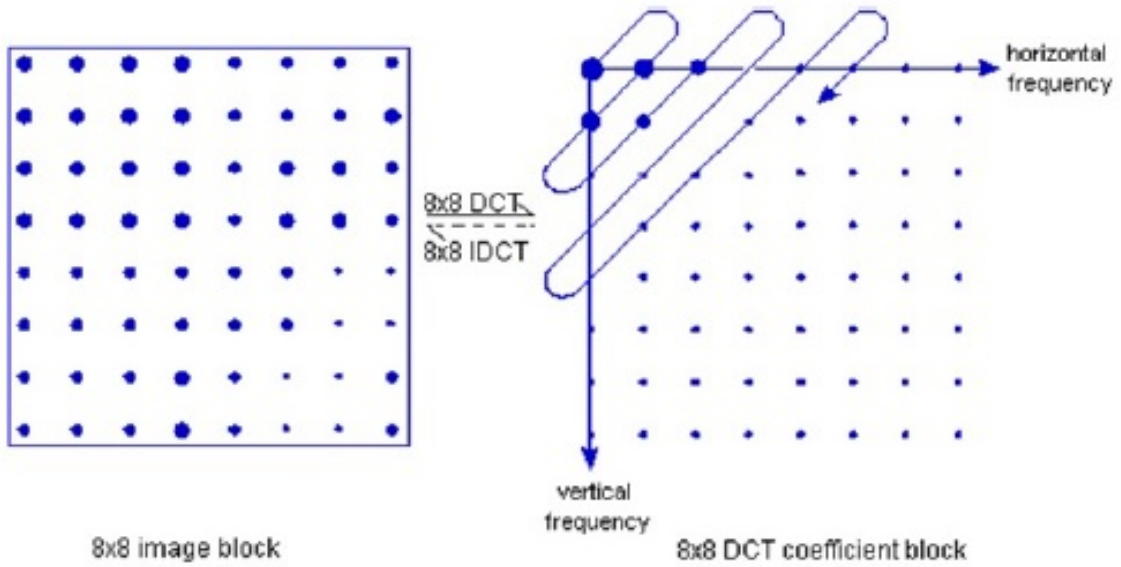
- The representation of the colors in the image is converted from RGB to YCbCr.
- The resolution of the chroma data is reduced, usually by downsampling.
- The image is split into blocks of 8×8 pixels, and for each block, each of the Y, Cb, and Cr data undergoes a Discrete Cosine Transform (DCT).
- The amplitudes of the frequency components are quantized, the high-frequency components are discarded.
- The resulting data for all 8×8 blocks is further compressed with a lossless algorithm, a variant of Huffman encoding.
- The decoding process reverses these steps.

JPEG Image Compression



RLE: Run-length encoding
VLC: Variable-length code

Zigzag Coding



MATLAB: Example

```
I = imread('cameraman.tif');
I = im2double(I);
T = dctmtx(8);
dct = @(block_struct) T * block_struct.data * T';
B = blockproc(I,[8 8],dct);
mask = [1  1  1  1  0  0  0  0
        1  1  1  0  0  0  0  0
        1  1  0  0  0  0  0  0
        1  0  0  0  0  0  0  0
        0  0  0  0  0  0  0  0
        0  0  0  0  0  0  0  0
        0  0  0  0  0  0  0  0
        0  0  0  0  0  0  0  0];
B2 = blockproc(B,[8 8],@(block_struct) mask .*
block_struct.data);
invdct = @(block_struct) T' * block_struct.data * T;
I2 = blockproc(B2,[8 8],invdct);
imshow(I), figure, imshow(I2)
```

website: https://au.mathworks.com/help/images/discrete-cosine-transform.html

MPEG Video Compression

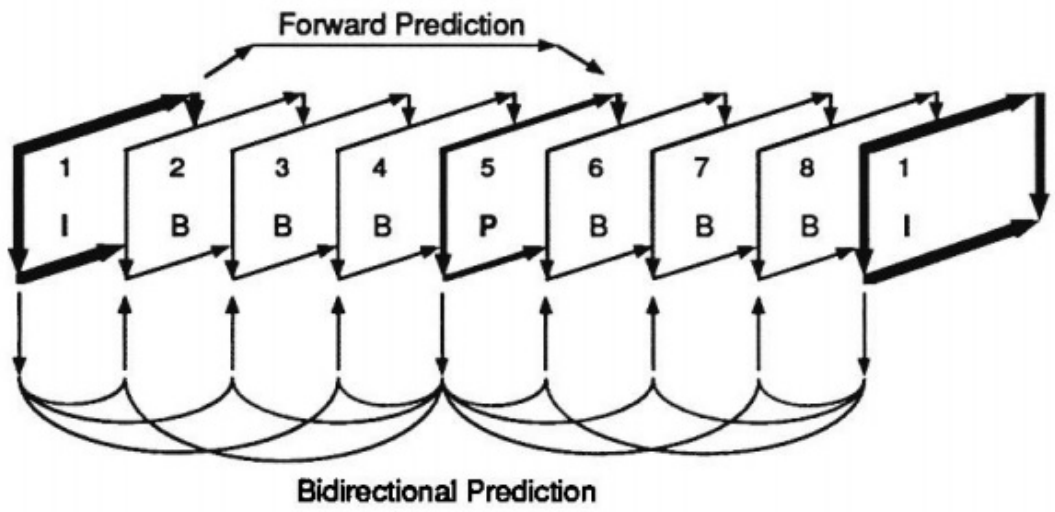
- MPEG stands for the Moving Picture Experts Group (1988, Canada).
- MPEG-1(1992): For VCD formats, up to 1.5 Mbits per second.
- MPEG-2(1994): For digital broadcast television, the bit rates between 1.5 and 15 Mbits per second.
- MPEG-4(1999, 23 Parts): Advanced Video Coding (AVC), object-oriented composite files, etc.
- MPEG-7(2001, 10 Parts): For multimedia content description, it uses XML to store metadata in order to tag particular events.
- MPEG-21(2002, 9 Parts): For sharing digital rights, permissions and restrictions of digital content.

H.26X Video Compression

- H.261 (1990): Designed for dual communication over ISDN lines, supports data rates 64Kbits per second. The algorithm is based on DCT and uses Intraframe (I-Frame coding) and Interframe (P-Frame coding) compression.
- H.264 (2003): Issued by the ITU-T VCEG together with MPEG. The ITU-T H.264 standard and the ISO/IEC MPEG-4 Part 10 or AVC (Advanced Video Coding) standard is jointly maintained so that they have identical technical content.
- H.265 (2013): High Efficiency Video Coding (HEVC), scale from 320×240 pixels up to 7680×4320 (8K).
- H.266 (2023): Versatile Video Coding (VVC) for streaming UHD content in 4K, 8K and even 16K on TV as well as High Dynamic Range (HDR) and 360° video.

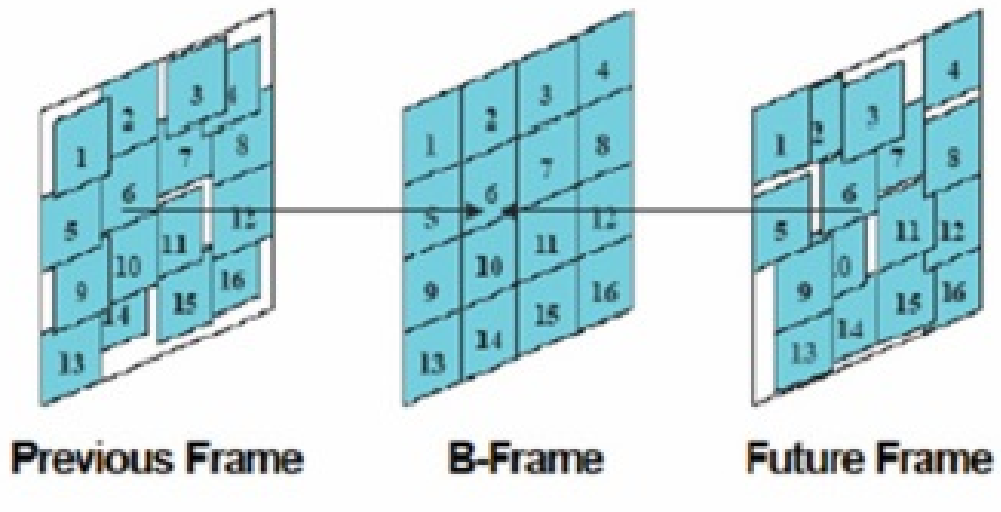
GOP: Group of Pictures

- *I-Frame*: Intra-coded picture or frame
- *B-Frame*: Bi-directionally predicted picture or frame
- *P-Frame*: Inter-coded picture or forwarded prediction frame
- *Motion Vector*: To present the position of a macro-block in another picture



Motion Vectors

- *I-Frame*: Intra-coded picture or frame
- *B-Frame*: Bi-directionally predicted picture or frame
- *P-Frame*: Inter-coded picture or forwarded prediction frame
- *Motion Vector*: To present the position of a macro-block in another picture



Questions?

Questions?

Which video compression standard is the most popular one at present?

1. MPEG-1
2. MPEG-4
3. MPEG-7
4. MPEG-21

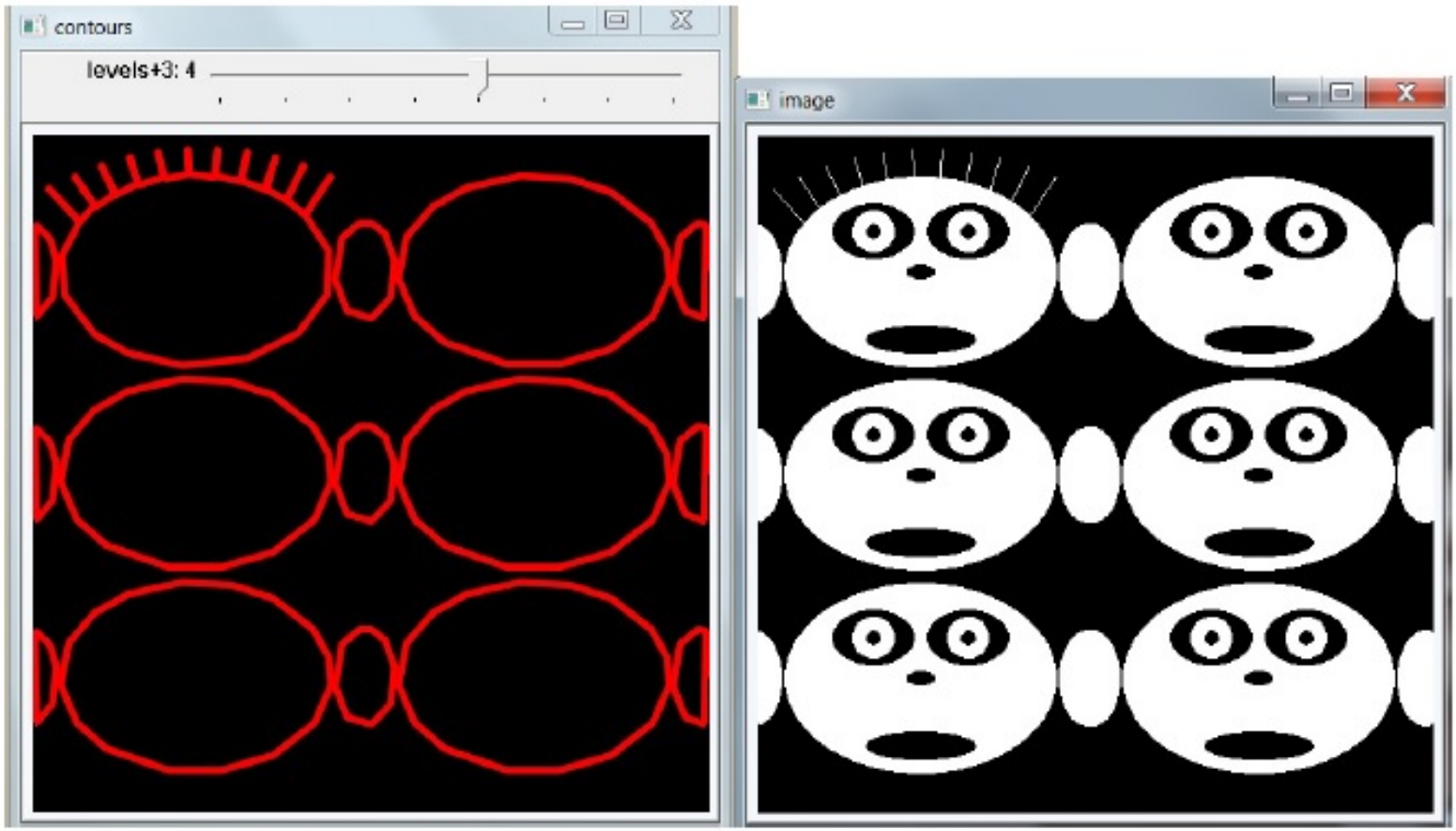
The right answer is:___

Questions?

OpenCV: Examples

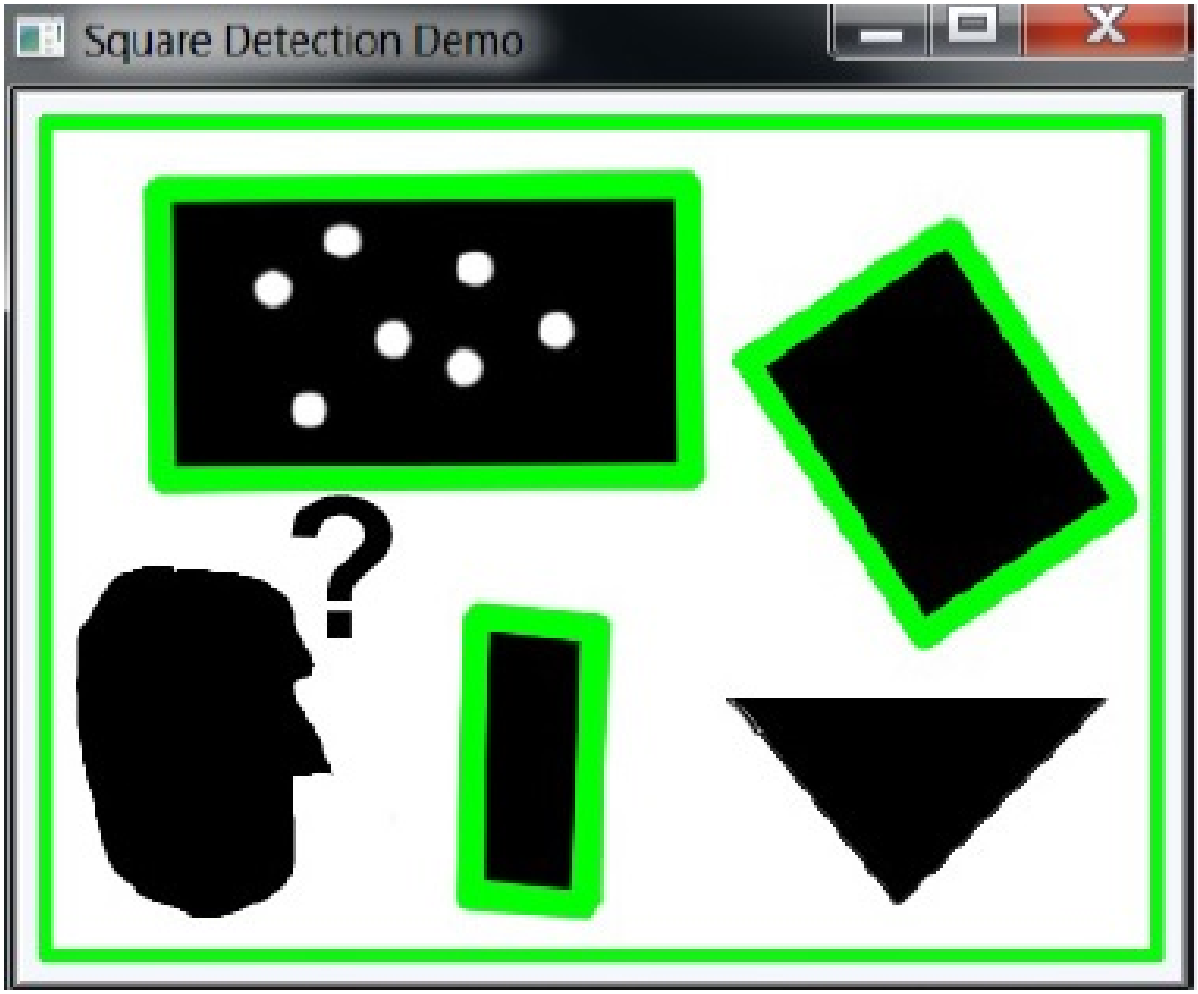
- Contours
- Edge extraction
- Hough lines
- k-means
- Morphology
- PCA: Principle component analysis
- Object segmentation
- Object detection
- Object tracking
- ...

OpenCV: Contour

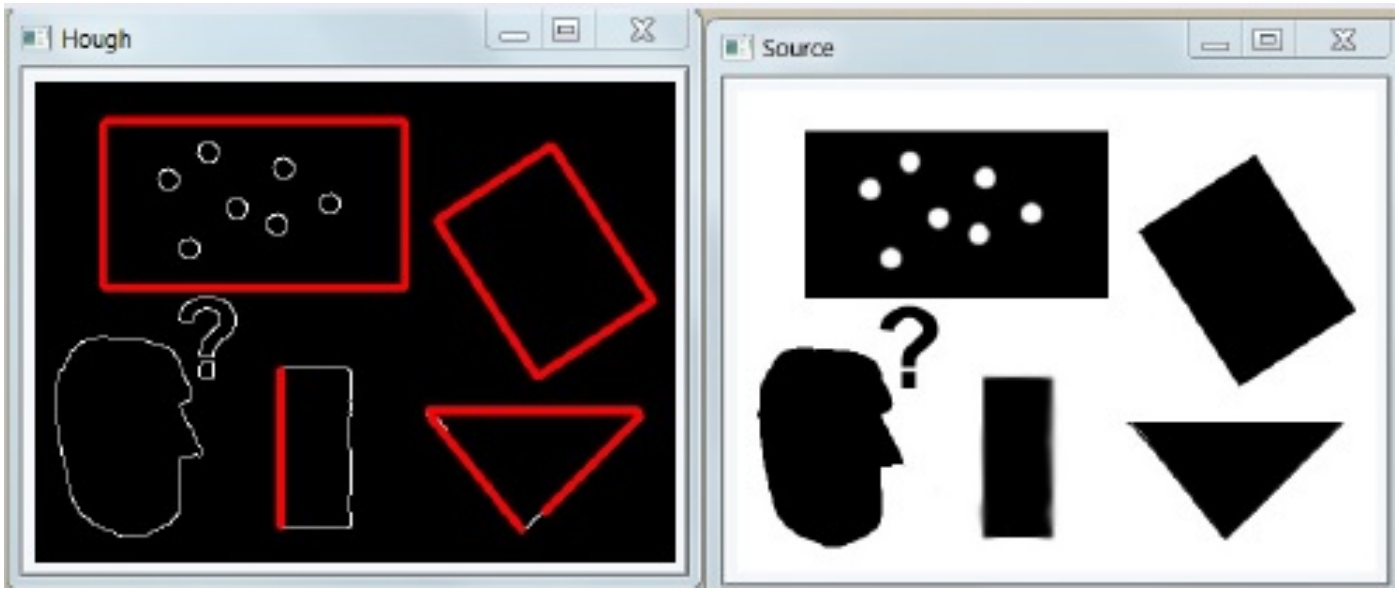


OpenCV: SIFT Matching

OpenCV: Square Detection



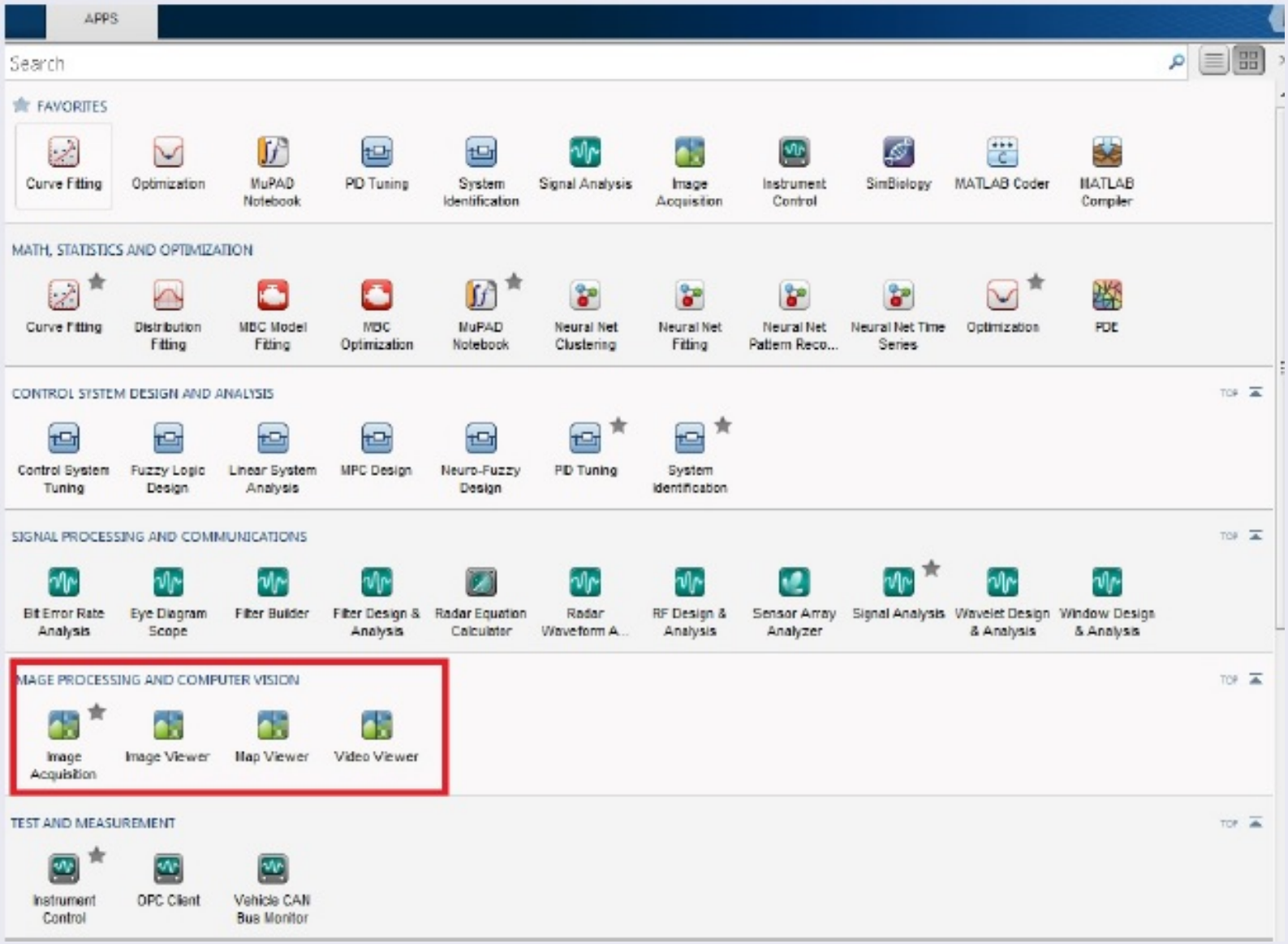
OpenCV: Hough Line Detection



MATLAB: Key Features

- Image analysis, including segmentation, morphology, statistics, and measurement
- Image enhancement, filtering, and deblurring
- Geometric transformations and intensity-based image registration methods
- Image transforms, including FFT, DCT, etc.
- Large image workflows, including block processing, tiling, and multiresolution display
- Visualization Apps, including image viewer and video viewer
- Multicore- and GPU-enabled functions, and C-code generation support

MATLAB: Toolboxes



Questions?

Questions?

What's the difference between OpenCV and MATLAB in digital image processing and computer vision?

1. OpenCV has the toolbox, MATLAB has not the toolbox.
2. MATLAB has the toolbox, OpenCV has not the toolbox.
3. MATLAB and OpenCV have the toolbox.
4. MATLAB and OpenCV have not the toolbox.

The right answer is:___

Questions?

Vision Intelligence I: Introduction

Learning Objectives

- To get organized for the course and to provide an overview of the core issues.
- Demonstrate understanding of the basic concepts, approaches, and algorithms used in vision intelligence.