# ***Computer Vision***

Course 1

## Online instruments

- ✓ **Webex – course, lab, article presentation**
- ✓ **Discord – lab, offline messages**
- ✓ **anca.ignat@info.uaic.ro, ignata@uaic.ro,**

  **ancai_fii@yahoo.ro**

  **https://edu.info.uaic.ro/computer-vision/**

# Evaluation

**1(base) + 3 (course test) + 3 (lab activity)+ 3 (article presentation)**

**$\geq$ 4.5**

**Course test (onsite!!!) – week 15 or 16**

**Lab activity – 6 projects**

**Lab evaluation (onsite!!!) – deadline week 12**

**Article presentation – weeks 13 and 14 (online)**

**At the end of the semester, send an archive with (code lab activity + emoji as .jpg file + article related documents) to ancai_fii@yahoo.ro**

## Course 1

# Bibliography


- **E.R. Davies, *Computer vision: principles, algorithms, applications, learning*, 5-th ed., Academic Press (2017)**
- **D.A.Forsyth & J. Ponce, *Computer vision: a modern approach*, Prentice Hall (2002)**
- **R. Klette, *Concise computer vision* (Vol. 233). London: Springer (2014)**
- **R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature (2022).**
- **R.C. Gonzales, R.E. Woods, *Digital Image Processing*, Pearson India, 2018, 4th ed.**
- **R. Shanmugamani, *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*, Packt Publishing Ltd. (2018).**
- **J. Brownlee, *Deep learning for computer vision: image classification, object detection, and face recognition in python*, Machine Learning Mastery (2019).**

## Course 1

- **Barbara J. Grosz (Harvard, USA):**

**“An AI and Computer Science Dilemma: Could I? Should I?“**

**(DIGHUM Lectures Program)**

- **face recognition ethic problem**

**Course 1**

"A very real concern for everyone involved in imaging, particularly in scientific and forensic fields, is the question of what constitutes proper and appropriate processing, and what constitutes unethical or even fraudulent manipulation. **The short answer is that anything that alters an image so as to create a false impression on the part of the viewer is wrong.** The problem with that answer is that it does not take into account the fact that different viewers will see different things in the image anyway, and that what constitutes a false impression for one person may not for another. The first rule is always to store a permanent copy of the original image along with relevant data on its acquisition. The second rule is to carefully document whatever steps are taken to process the image and generally to report those steps when the processed image is published."

John C. Russ, F. Brent Neal – The Image Processing Handbook, 7-th ed., CRC Press, 2016

## Course 1

# Meet Lena!

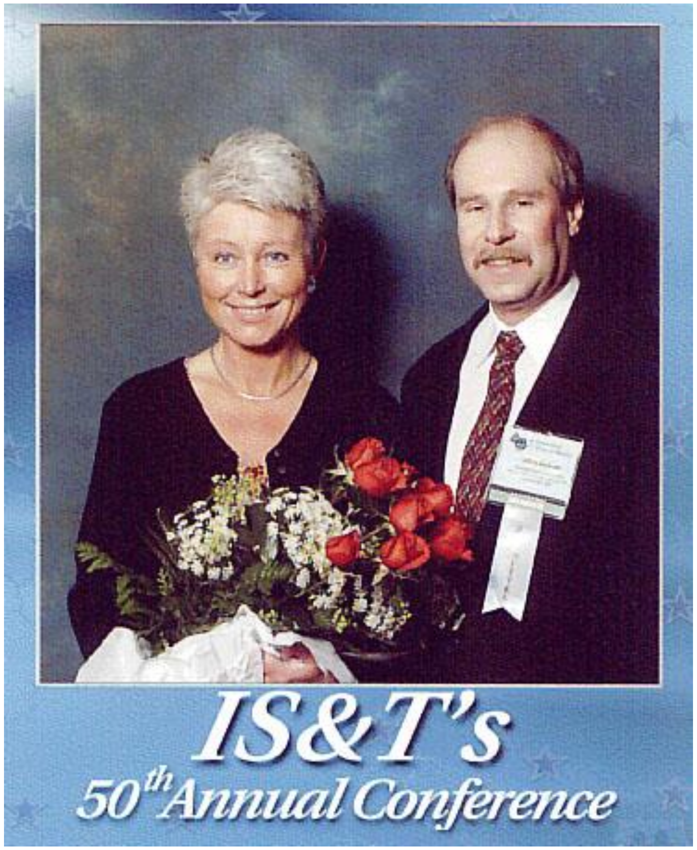
## The First Lady of the Internet

https://www.g4media.ro/o-fotografie-playboy-din-1972-interzisa-de-o-comunitate-stiintifica-fotografia-cu-modelul-lena-forsen-a-devenit-standardul-jpeg-e-considerata-acum-o-imagine-sexista.html

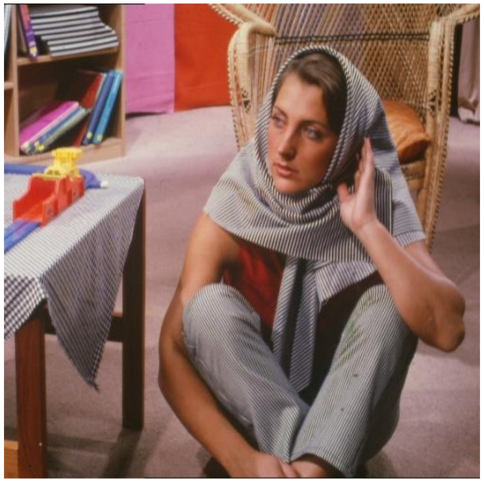
## Course 1

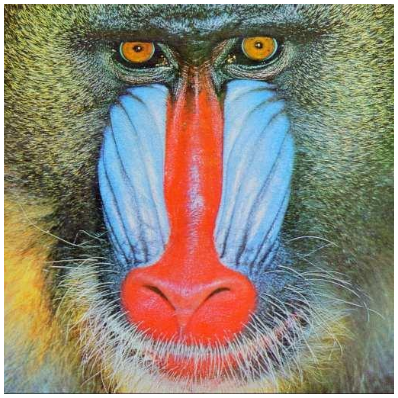
**Lenna Soderberg (Sjööblom) and  Jeff Seideman taken in May 1997**

**Imaging Science & Technology   Conference**



# Course 1





# Course 1

## What is Computer Vision?

**image processing , image analysis , computer vision ?**

Image processing = discipline in which both the input and the output of a process are images

Computer Vision = use computer to emulate human vision (AI) – learning, making inferences and take actions based on visual inputs

Image analysis (image understanding) = segmentation, partitioning images into regions or objects
(link between image processing and computer vision)

## Course 1

Distinction between image processing , image analysis , computer vision:

low-level, mid-level, high-level processes

Low-level processes:

image preprocessing to reduce noise, contrast enhancement,
image sharpening; both input and output are images

Mid-level processes:

- segmentation, partitioning images into regions or objects,
- description of the objects for computer processing,
- classification/recognition of individual objects;
- inputs are generally images, outputs are attributes extracted from the
- input image (e.g. edges, contours, identity of individual objects)

High-level processes:

- “making sense” of a set of recognized objects;
- performing the cognitive functions associated with vision

According to Gonzales + Woods (Digital Image Processing)

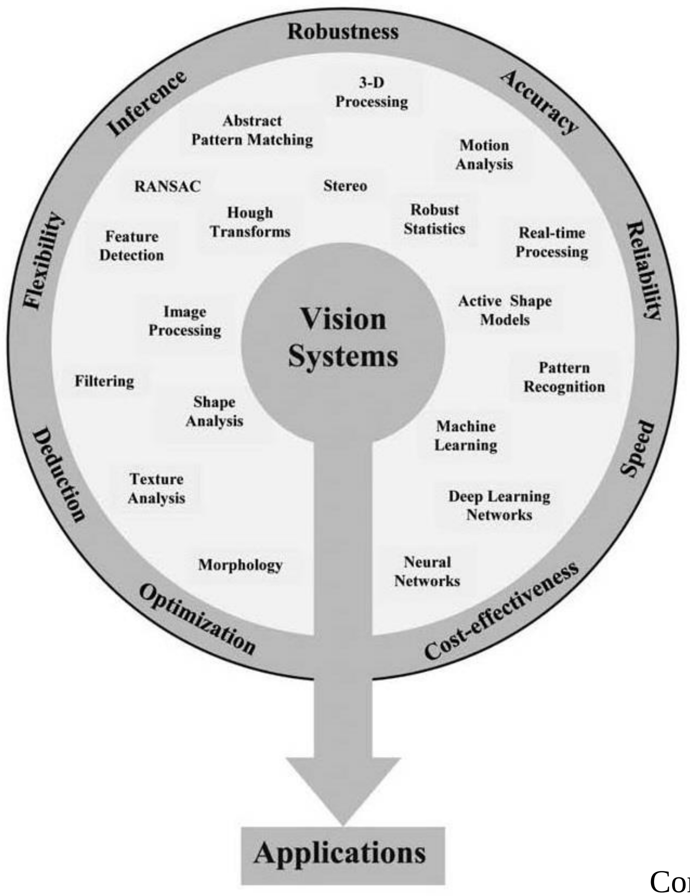
Digital Image Processing = low-level processes + mid-level processes

Computer Vision = low-level processes + mid-level processes +

+ high-level processes

# Course 1

INFLUENCES IMPINGING UPON INTEGRATED VISION SYSTEM DESIGN




Computer Vision Principles, Algorithms, Applications, Learning (E.R. Davies)

## Course 1

**https://www.ibm.com/topics/computer-vision**

**Computer vision** is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs — and take actions or make recommendations based on that information. If AI enables computers to think, computer vision enables them to see, observe and understand.

Computer vision works much the same as human vision, except humans have a head start. Human sight has the advantage of lifetimes of context to train how to tell objects apart, how far away they are, whether they are moving and whether there is something wrong in an image.

Computer vision trains machines to perform these functions, but it has to do it in much less time with cameras, data and algorithms rather than retinas, optic nerves and a visual cortex. Because a system trained to inspect products or watch a production asset can analyze thousands of products or processes a minute, noticing imperceptible defects or issues, it can quickly surpass human capabilities.

## Course 1

# Wiki

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do.

Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, e.g. in the forms of decisions. Understanding in this context means the transformation of visual images (the input of the retina) into descriptions of the world that make sense to thought processes and can elicit appropriate action. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory.

The scientific discipline of computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, multi-dimensional data from a 3D scanner, or medical scanning device. The technological discipline of computer vision seeks to apply its theories and models to the construction of computer vision systems.

Sub-domains of computer vision include scene reconstruction, object detection, event detection, video tracking, object recognition, 3D pose estimation, learning, indexing, motion estimation, visual serving, 3D scene modeling, and image restoration.

Jason Brownlee - Deep Learning for Computer Vision - Image Classification, Object Detection, and Face Recognition in Python

## 1.3 What Is Computer Vision

Computer vision is a field of study focused on the problem of helping computers to see.

> At an abstract level, the goal of computer vision problems is to use the observed image data to infer something about the world.
>
> — Page 83, *Computer Vision: Models, Learning, and Inference*, 2012.

It is a multidisciplinary field that could broadly be called a subfield of artificial intelligence and machine learning, which may involve the use of specialized methods and make use of general learning algorithms.

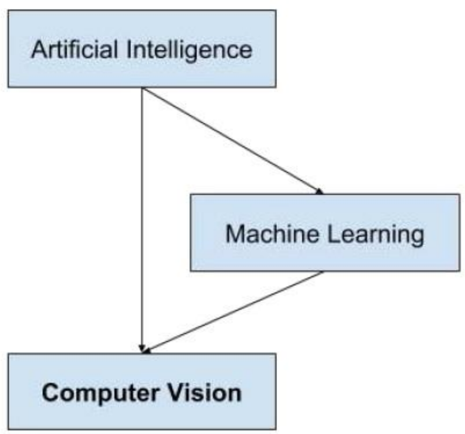



Figure 1.1: Overview of the Relationship of Artificial Intelligence and Computer Vision.

As a multidisciplinary area of study, it can look messy, with techniques borrowed and reused from a range of disparate engineering and computer science fields. One particular problem in vision may be easily addressed with a hand-crafted statistical method, whereas another may require a large and complex ensemble of generalized machine learning algorithms.

> Computer vision as a field is an intellectual frontier. Like any frontier, it is exciting and disorganized, and there is often no reliable authority to appeal to. Many useful ideas have no theoretical grounding, and some theories are useless in practice; developed areas are widely scattered, and often one looks completely inaccessible from the other.
>
> — Page xvii, *Computer Vision: A Modern Approach*, 2002.

The goal of computer vision is to understand the content of digital images. Typically, this involves developing methods that attempt to reproduce the capability of human vision. Understanding the content of digital images may involve extracting a description from the image, which may be an object, a text description, a three-dimensional model, and so on.

> Computer vision is the automated extraction of information from images. Information can mean anything from 3D models, camera position, object detection and recognition to grouping and searching image content.
>
> — Page ix, *Programming Computer Vision with Python*, 2012.

## Course 1

### What is a Digital Image?

$$f : D \to \mathbb{R} \ / \ \mathbb{R}^{3/4}$$

***f(x,y)*** = intensity, gray level, color of the image at spatial point ***(x,y)***

*x*, *y*, *f(x,y)* – finite, discrete quantities → digital image

Digital Image Processing = processing digital images by means of a digital computer

A digital image is composed of a finite number of elements (location, value of intensity):

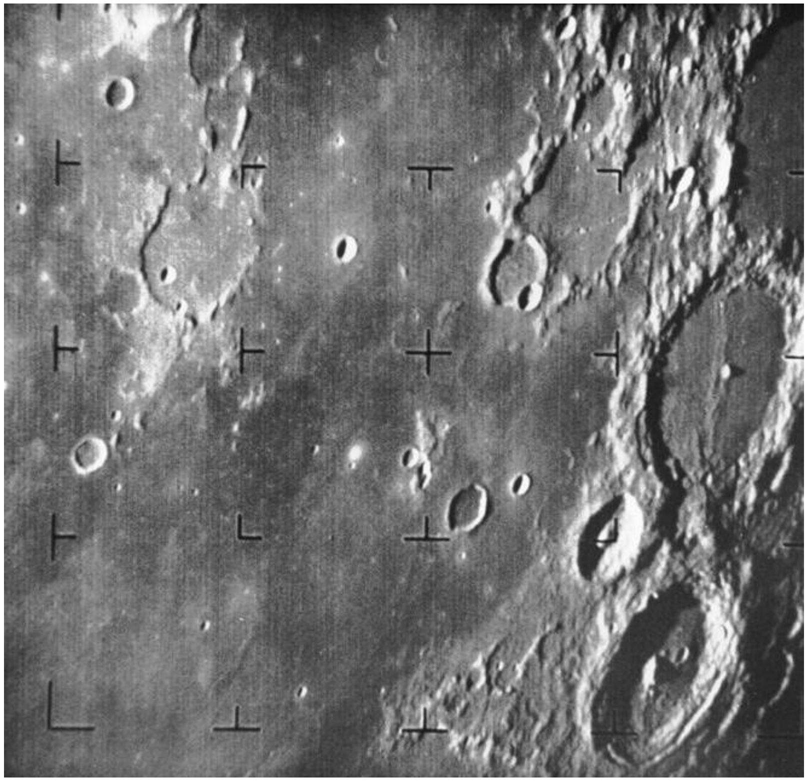
$$(x_i, y_j, f_{ij})$$

These elements are called ***picture elements***, ***image elements***, ***pels***, ***pixels***

## Course 1

## The Beginning

1964, Jet Propulsion Laboratory (Pasadena, California) processed pictures of the moon transmitted by Ranger 7 (corrected image distortions)



The first picture of the moon by a U.S. spacecraft. *Ranger 7* took this image July 31, 1964, about 17 minutes before impacting the lunar surface.
(Courtesy of NASA)

1960-1970 – image processing techniques were used on medical images,

remote Earth resources observations, astronomy

1970s – invention of CAT (computerized axial tomography)

CAT is a process in which a ring of detectors encircles an object (patient), and an X-ray source, concentric with the detector ring, rotates about the object.

The X-ray passes through the patient and the resulted waves are collected at the opposite end by the detectors. As the source rotates the procedure is repeated.

Tomography consists of algorithms that use the sense data to construct an image that represents a “slice” through the object. Motion of the object in a direction perpendicular to the ring of detectors produces a set of “slices” which can be assembled in a 3D information of the inside of the object.

◊ geographers use CV to study pollution patterns from aerial and satellite imagery

◊ archeology – DIP+CV allowed restoring blurred pictures that recorded rare artifacts lost or damaged after being photographed

◊ physics – enhance images of experiments (high-energy plasmas, electron microscopy)

◊ astronomy, biology, nuclear medicine, law enforcement, industry

◊ automatic character recognition, industrial machine vision for product assembly and inspection, military recognizance, automatic processing of fingerprints, automatic processing of aerial and satellite imagery for weather prediction, Internet

## *Examples of CV Applications*

Images can be classified according to the energy sources that reveals the scene (visual, X-ray, …)

Energy sources for images:

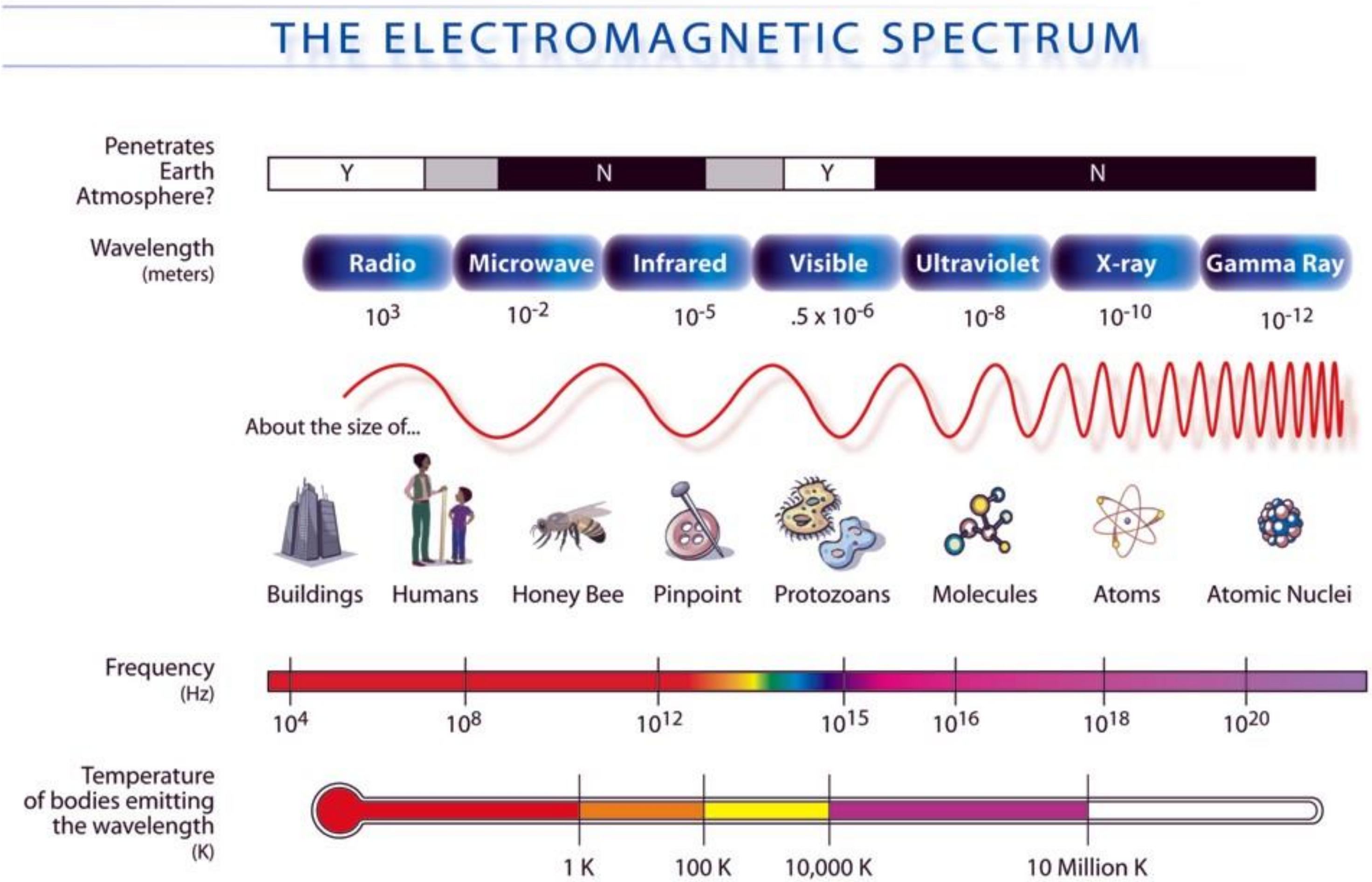
electromagnetic energy spectrum,

acoustic,

ultrasonic,

electronic,

computer-generated


THE ELECTROMAGNETIC SPECTRUM
Penetrates Earth Atmosphere?
Y
N
Y
N
Wavelength (meters)
Radio
Microwave
Infrared
Visible
Ultraviolet
X-ray
Gamma Ray
$10^3$
$10^{-2}$
$10^{-5}$
$.5 \times 10^{-6}$
$10^{-8}$
$10^{-10}$
$10^{-12}$
About the size of...
Buildings
Humans
Honey Bee
Pinpoint
Protozoans
Molecules
Atoms
Atomic Nuclei
Frequency (Hz)
$10^4$
$10^8$
$10^{12}$
$10^{15}$
$10^{16}$
$10^{18}$
$10^{20}$
Temperature of bodies emitting the wavelength (K)
1 K
100 K
10,000 K
10 Million K

## Gamma-Ray Imaging

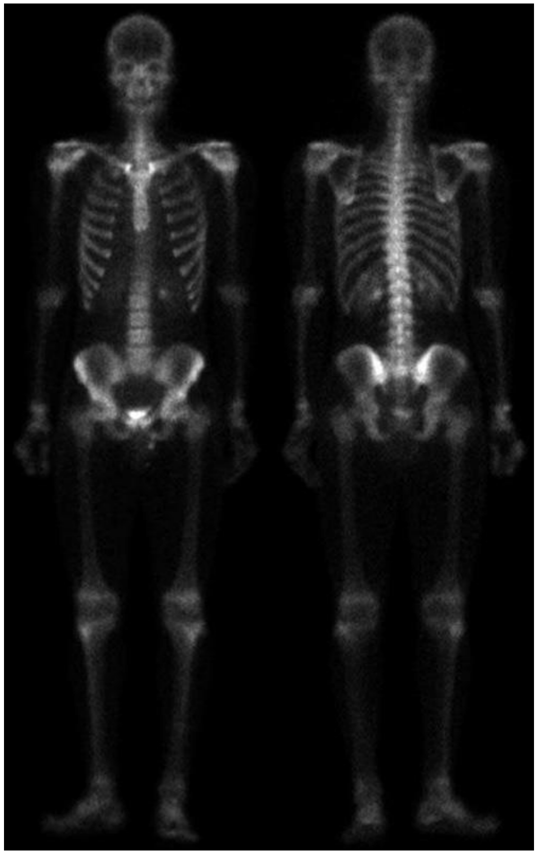
Nuclear medicine, astronomical observations

Nuclear medicine

The approach is to inject a patient with a radioactive isotope that emits gamma rays as it decays. Images are produced from the emissions collected by gamma-ray detectors.
Images of this sort are used to locate sites of bone pathology (infections, tumors)
PET (positron emission tomography) – the patient is given a radioactive isotope that emits positrons as it decays

# Examples of gamma-ray imaging

Bone scan



PET image

## Course 1

### X-ray imaging

Medical diagnostic, industry, astronomy

An X-ray tube is a vacuum tube with a cathode and an anode. The cathode is heated, causing free electrons to be released. The electrons flows at high speed to the positively charged anode. When the electrons strike a nucleus, energy is released in the form of an X-ray radiation. The energy (penetrating power) of the X-rays is controlled by a voltage applied across the anode, and by a current applied to the filament in the cathode. The intensity of the X-rays is modified by absorption as they pass through the patient and the resulting energy falling develops it much in the same way that light develops photographic film.

## Course 1

Angiography = contrast-enhancement radiography

Angiograms = images of blood vessels

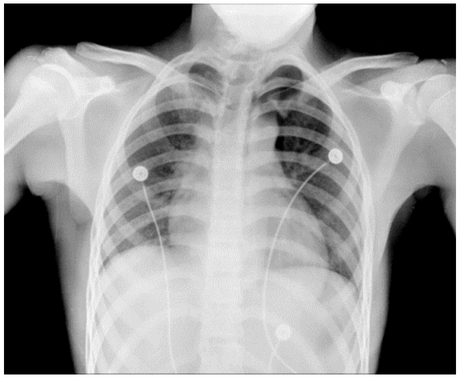
A catheter is inserted into an artery or vein in the groin. The catheter is threaded into the blood vessel and guided to the area to be studied. When it reaches the area to be studied, an X-ray contrast medium is injected through the catheter. This enhances contrast of the blood vessels and enables radiologist to see any irregularities or blockages.

X-rays are used in CAT (computerized axial tomography)

X-rays used in industrial processes (examine circuit boards for flows in manufacturing)

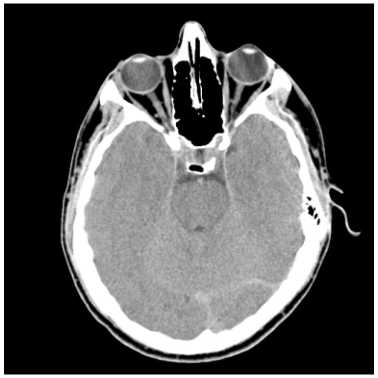
Industrial CAT scans are useful when the parts can be penetrated by X-rays

**Examples of X-ray images**

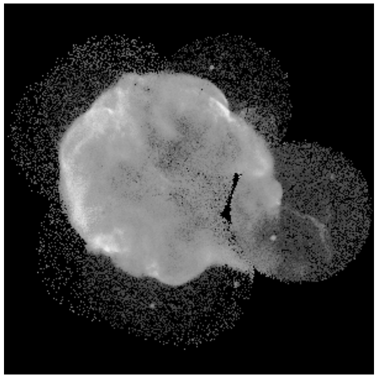


Chest X-ray

Aortic angiogram



Head CT



Cygnus Loop

**Course 1**

## Imaging in the Ultraviolet Band

→ industrial inspection, microscopy, biological imaging, astronomical observations

Ultraviolet light is used in fluorescence microscopy. Ultraviolet light is not visible to human eye but when a photon of ultraviolet radiation collides with an electron in an atom of a fluorescent material it elevates the electron to a higher energy level. After that the electron relaxes to a lower level and emits light in the form of a lower-energy photon in the visible (red) light region.

Fluorescence = emission of light by a substance that has absorbed light or other electromagnetic radiation of a different wavelength

Fluorescence microscope = uses an excitation light to irradiate a prepared specimen and then it separates the much weaker radiating fluorescent light from the brighter excitation light.

**Course 1**

## Imaging in the Visible and Infrared Bands

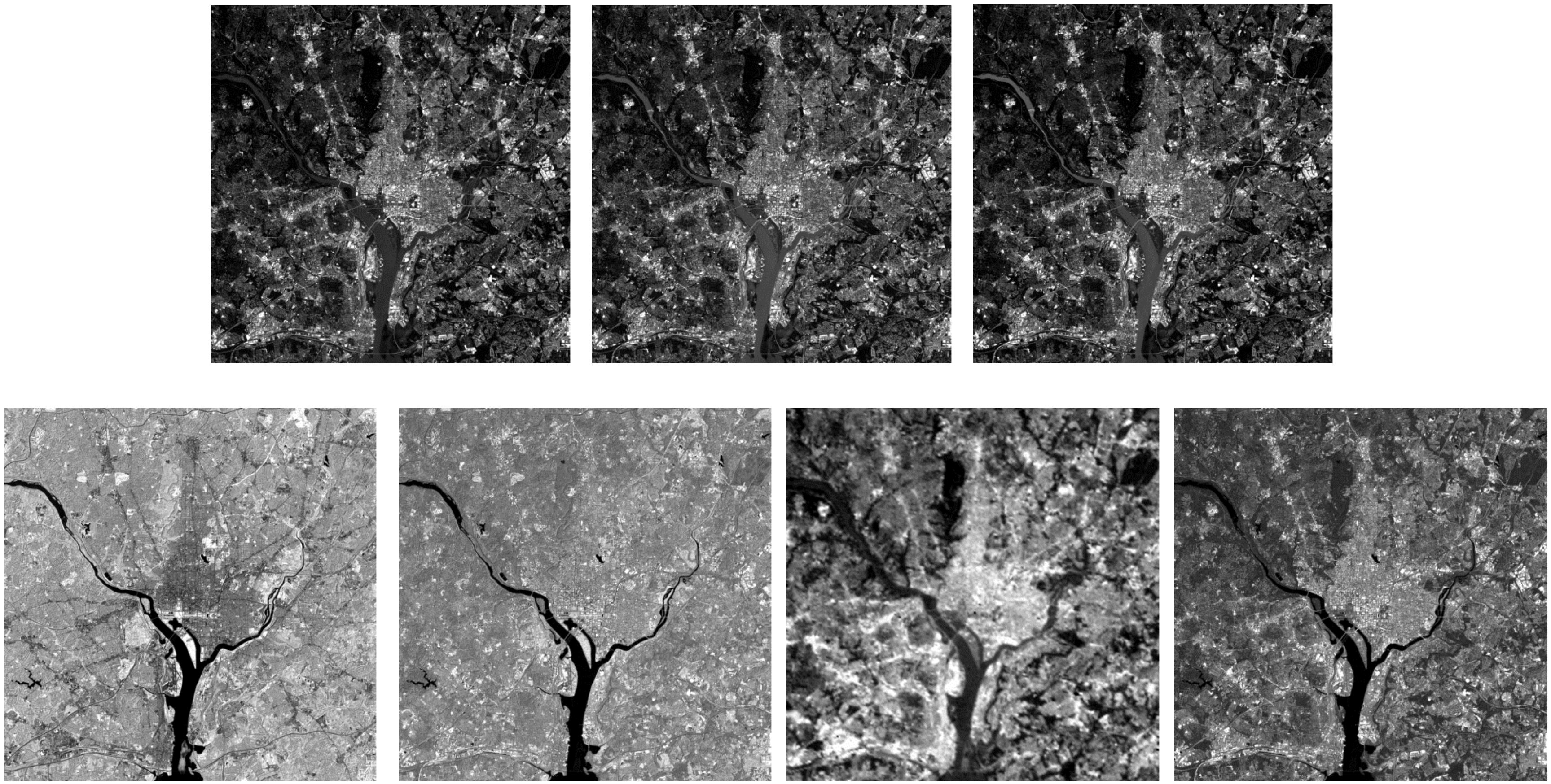
→ Light microscopy, astronomy, remote sensing, industry, law enforcement

LANDSAT satellite – obtained and transmitted images of the Earth from space
for purpose of monitoring environmental conditions on the planet

Weather observations and prediction produce major applications of multispectral image from satellites

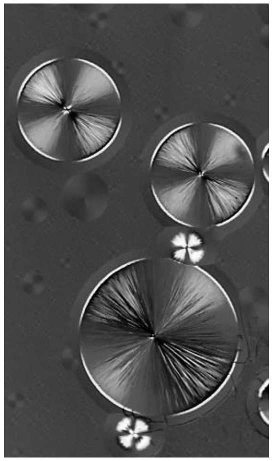
| Band No. | Name | Wavelength (μm) | Characteristics and Uses |
|---|---|---|---|
| 1 | Visible blue | 0.45–0.52 | Maximum water penetration |
| 2 | Visible green | 0.52–0.60 | Good for measuring plant vigor |
| 3 | Visible red | 0.63–0.69 | Vegetation discrimination |
| 4 | Near infrared | 0.76–0.90 | Biomass and shoreline mapping |
| 5 | Middle infrared | 1.55–1.75 | Moisture content of soil and vegetation |
| 6 | Thermal infrared | 10.4–12.5 | Soil moisture; thermal mapping |
| 7 | Middle infrared | 2.08–2.35 | Mineral mapping |

Thematic bands in NASA’s LANDSAT satellite

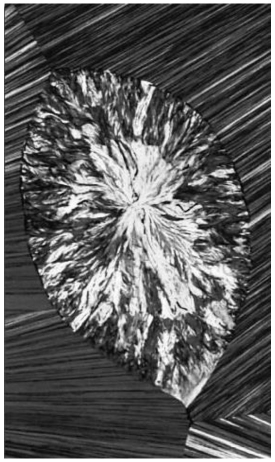


Satellite images of Washington D.C. area in spectral bands of the above table

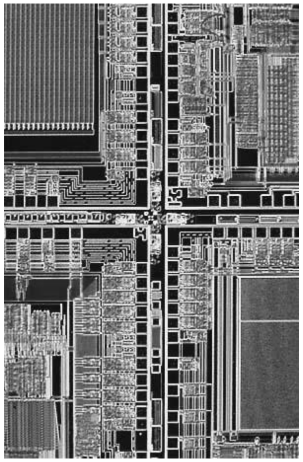
## Examples of light microscopy images



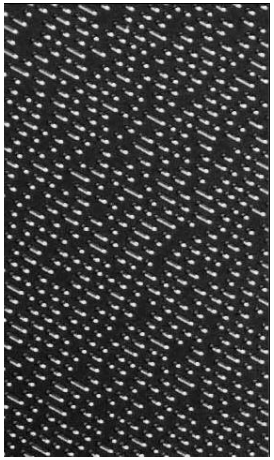
Taxol
(anticancer agent)
magnified 250X



Cholesterol
(40X)

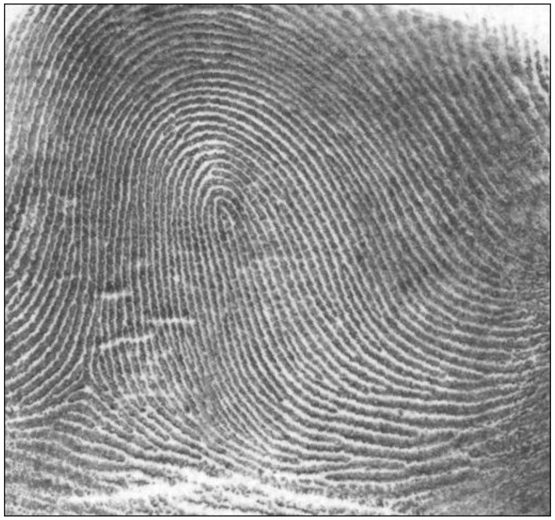


Microprocessor
(60X)



Surface of
audio CD
(1750X)

Organic
superconductor
(450X)




BANCA NAȚIONALĂ A ROMÂNIEI
GUVERNATOR
CASIER CENTRAL
1 LEU
UN LEU


BG
EH.0591 BM


IS 53 DEG
IF 20 MRC
IS 05 ZRC
IS 06 KRB
IS 09 DDO
B 65 ZAA
IS 28 CDC
B 98 XYN
IS 06 DCL
IS 08 KXU
IS 08 JMM
SB 06 PZM
IS 51 BNV
BT 27 MCA
B 39 YYD
BT 77 PAL

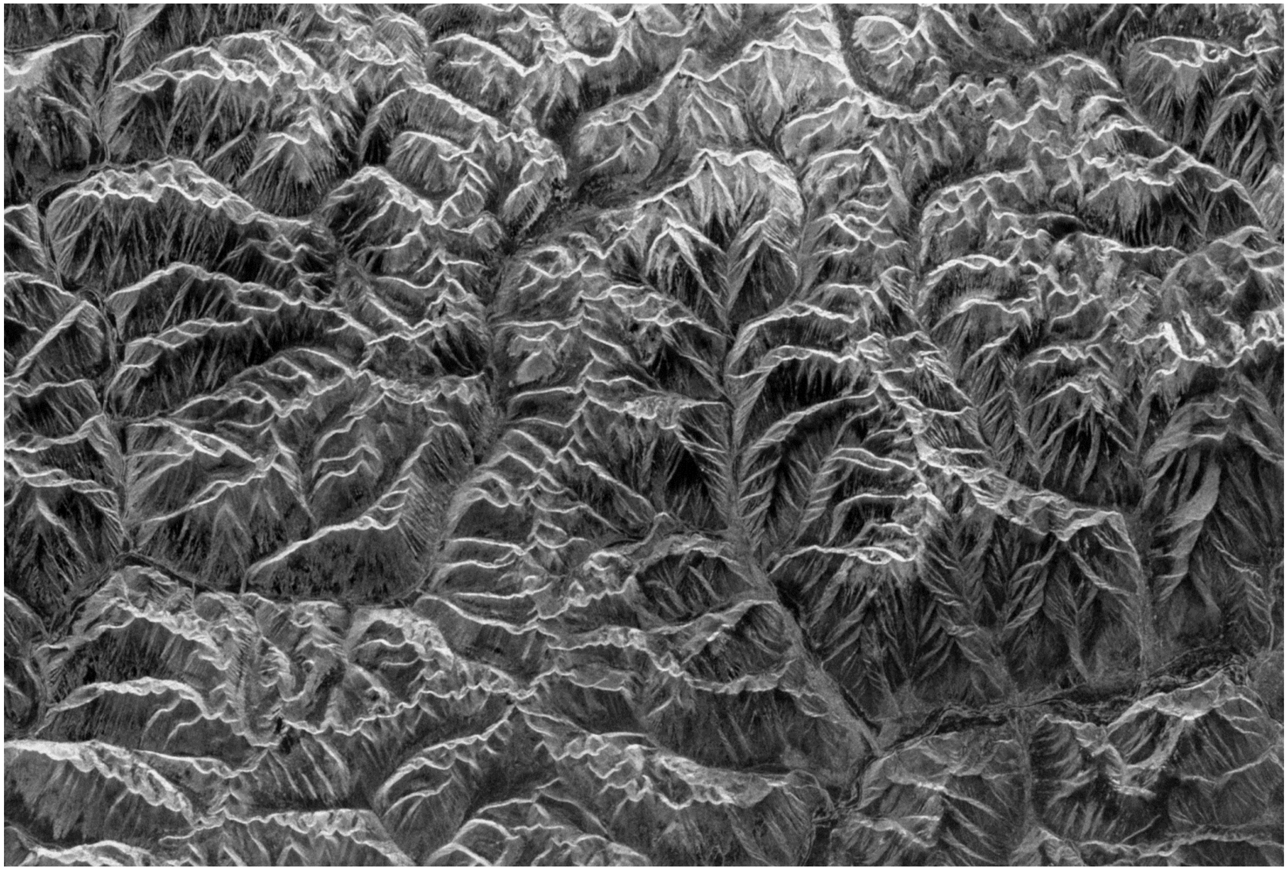
## Imaging in the Microwave Band

The dominant application of imaging in the microwave band – radar

- radar has the ability to collect data over virtually any region at any time, regardless of weather or ambient light conditions
- some radar waves can penetrate clouds, under certain conditions can penetrate vegetation, ice, dry sand
- sometimes radar is the only way to explore inaccessible regions of the Earth's surface

An imaging radar works like a flash camera: it provides its own illumination (microwave pulses) to light an area on the ground and take a snapshot image. Instead of a camera lens, a radar uses an antenna and a digital device to record the images. In a radar image one can see only the microwave energy that was reflected back toward the radar antenna.

**Course 1**



Spaceborne radar image of mountains in southeast Tibet

## Imaging in the Radio Band

→ medicine, astronomy

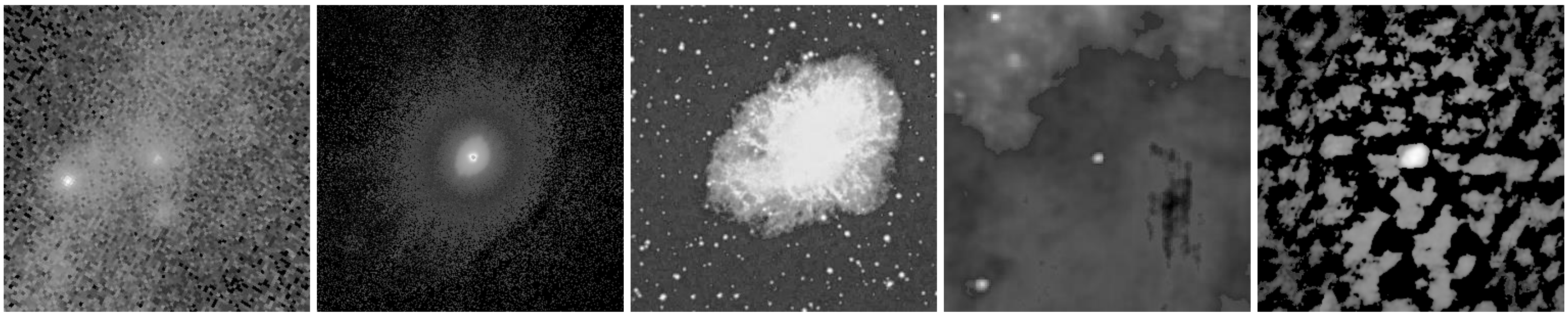
MRI = **M**agnetic **R**esonance **I**maging

This technique places the patient in a powerful magnet and passes short pulses of radio waves through his or her body. Each pulse causes a responding pulse of radio waves to be emitted by the patient tissues.

The location from which these signals originate and their strength are determined by a computer, which produces a 2D picture of a section of the patient.

MRI images of a human knee (left) and spine (right)

Images of the Crab Pulsar covering the electromagnetic spectrum



Gamma X-ray Optical Infrared Radio

## Other Imaging Modalities

→ acoustic imaging, electron microscopy, synthetic (computer-generated) imaging

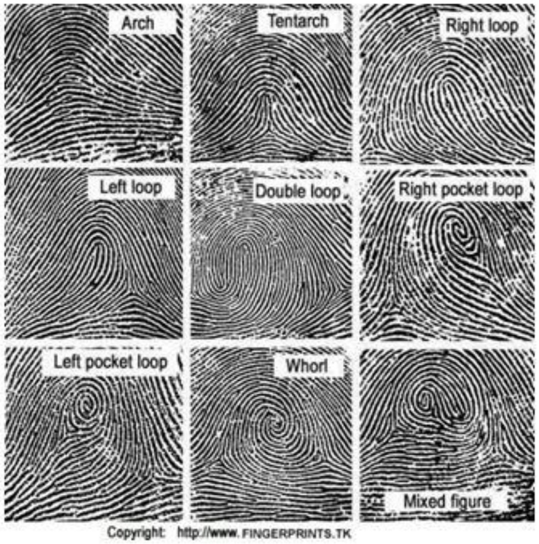
Imaging using sound → geological explorations, industry, medicine,
mineral and oil exploration

For image acquisition over land one of the main approaches is to use a large truck a large flat steel plate. The plate is pressed on the ground by the truck and the truck is vibrated through a frequency spectrum up to 100 Hz. The strength and the speed of the returning sound waves are determined by the composition of the Earth below the surface. These are analyzed by a computer and images are generated from the resulting analysis.

# Course 1

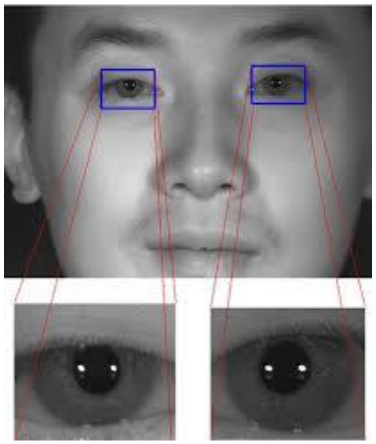
## Course 1

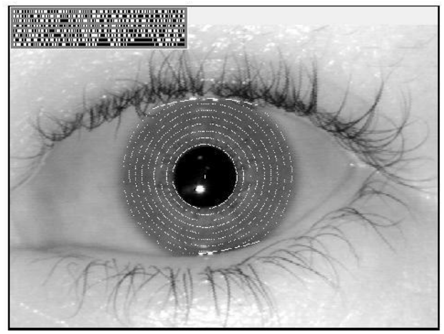
# Biometry – fingerprint

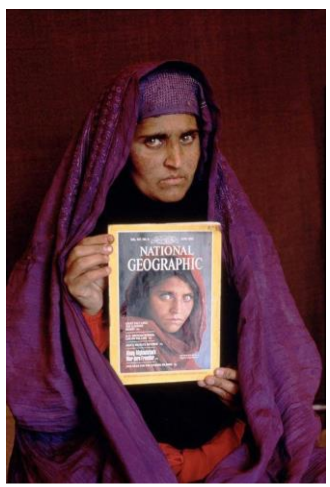




**Course 1**

# Biometry – iris

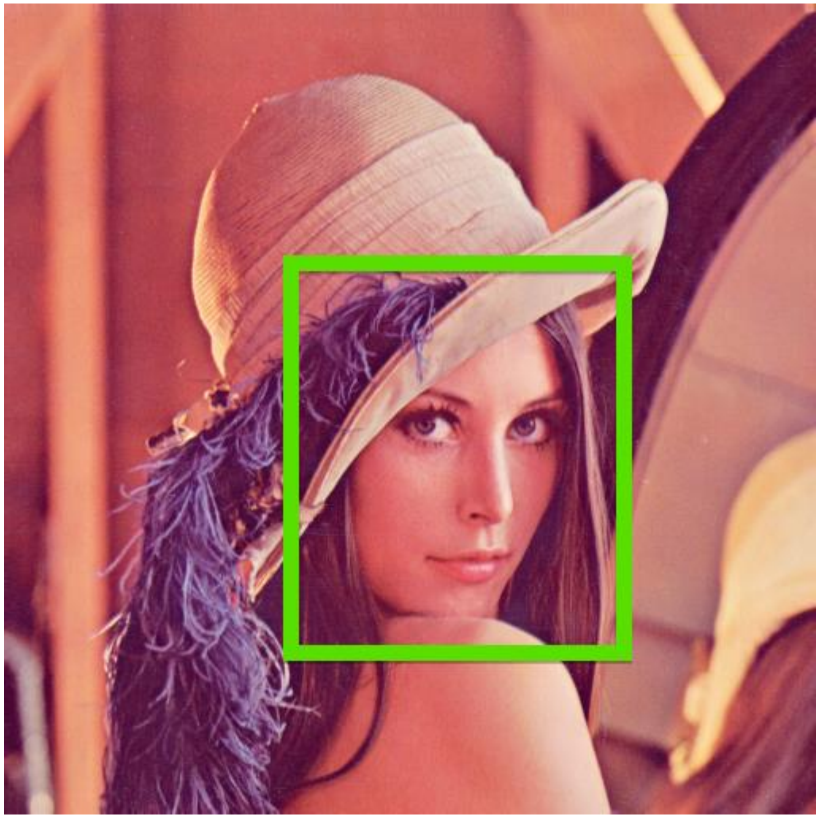







**Course 1**

# Face detection and recognition

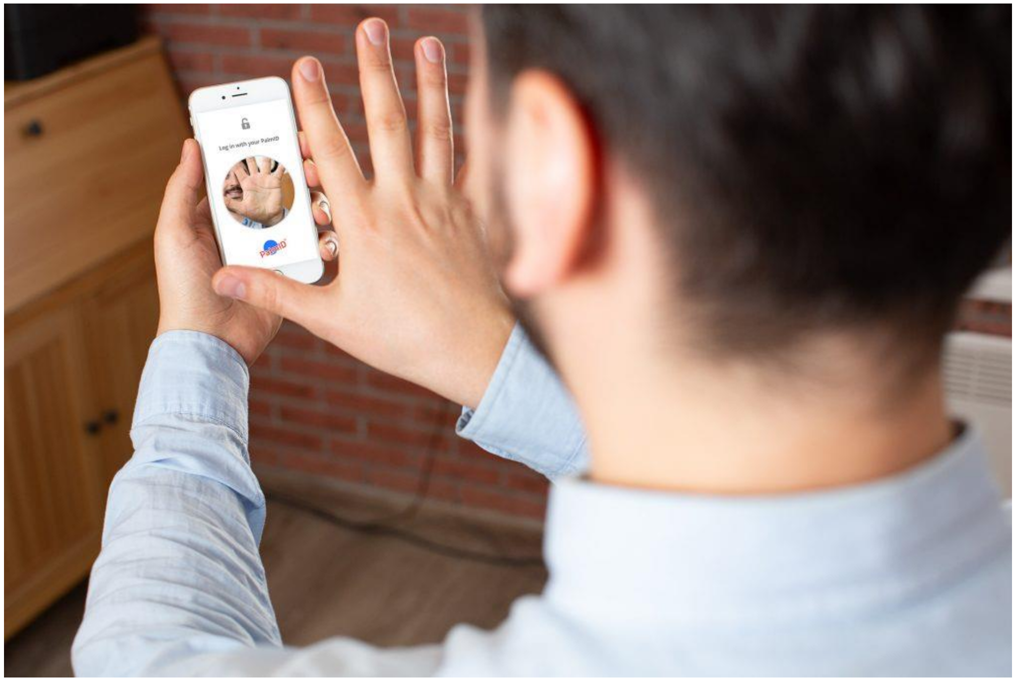


**Course 1**

# Gender identification

**Course 1**

# Palmprint authentication



**Course 1**

# Object detection/tracking

# Hawk-Eye Systems

# Deepfake

https://youtu.be/oxXpB9pSETo?si=JIJcG8IXccZq615X

https://youtu.be/iyiOVUbsPcM?si=Uj9WM200GT-VXlLu

## Fundamental Steps in Image Processing

- methods whose input and output are images
- methods whose inputs are images but whose outputs are attributes extracted from those images

## Course 1

### Outputs are images

- image acquisition
- image filtering and enhancement
- image restoration
- color image processing
- wavelets and multiresolution processing
- compression
- morphological processing

**Outputs are attributes**

- morphological processing
- segmentation
- representation and description
- object recognition

**Image acquisition -** may involve preprocessing such as scaling

**Image enhancement**

- manipulating an image so that the result is more suitable than the original for a specific operation
- enhancement is problem oriented
- there is no general ‚theory' of image enhancement
- enhancement use subjective methods for image improvement
- enhancement is based on human subjective preferences regarding what is a „good" enhancement result

**Image restoration**

- improving the appearance of an image
- restoration is objective - the techniques for restoration are based on mathematical or probabilistic models of image degradation

**Color image processing**

- fundamental concept in color models
- basic color processing in a digital domain

## Course 1

### Wavelets and multiresolution processing

- representing images in various degree of resolution

**Compression**

- reducing the storage required to save an image or the bandwidth required to transmit it

**Morphological processing**

- tools for extracting image components that are useful in the representation and description of shape
- a transition from processes that output images to processes that output image attributes

**Segmentation**

- partitioning an image into its constituents parts or objects
- autonomous segmentation is one of the most difficult tasks of DIP
- the more accurate the segmentation , the more likely recognition is to succeed

**Representation and description** (almost always follows segmentation)

- segmentation produces either the boundary of a region or all the points in the region itself
- converting the data produced by segmentation to a form suitable for computer processing

- boundary representation: the focus is on external shape characteristics such as corners or inflections
- complete region: the focus is on internal properties such as texture or skeletal shape
- description is also called feature extraction – extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another

**Object recognition** (Machine Learning techniques)

- the process of assigning a label (e.g. „vehicle”) to an object based on its descriptors

**Knowledge database**

# Course 1

## Course 1

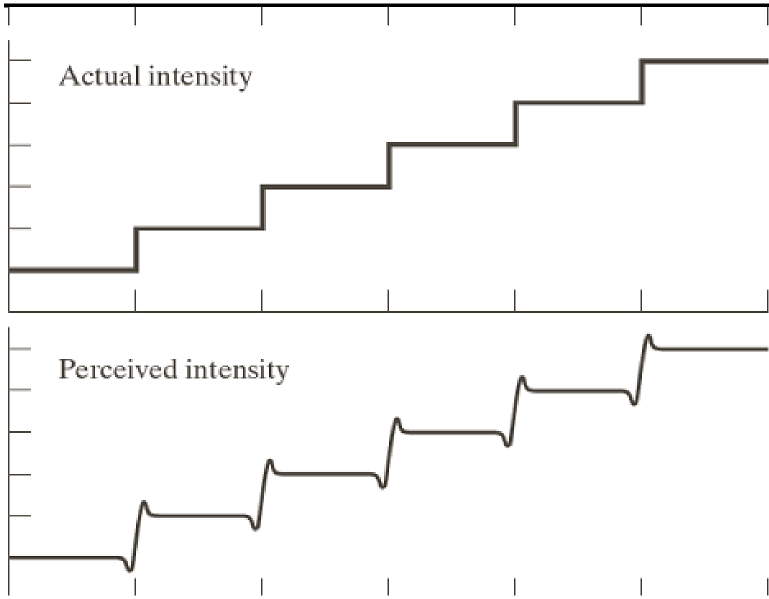
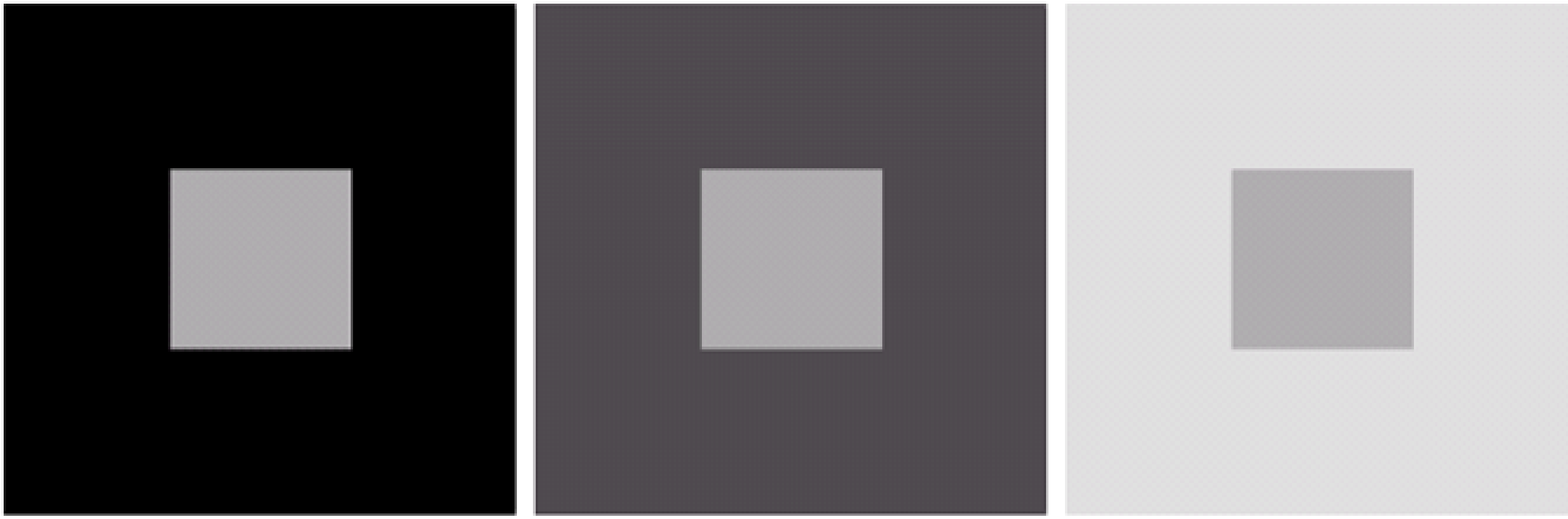


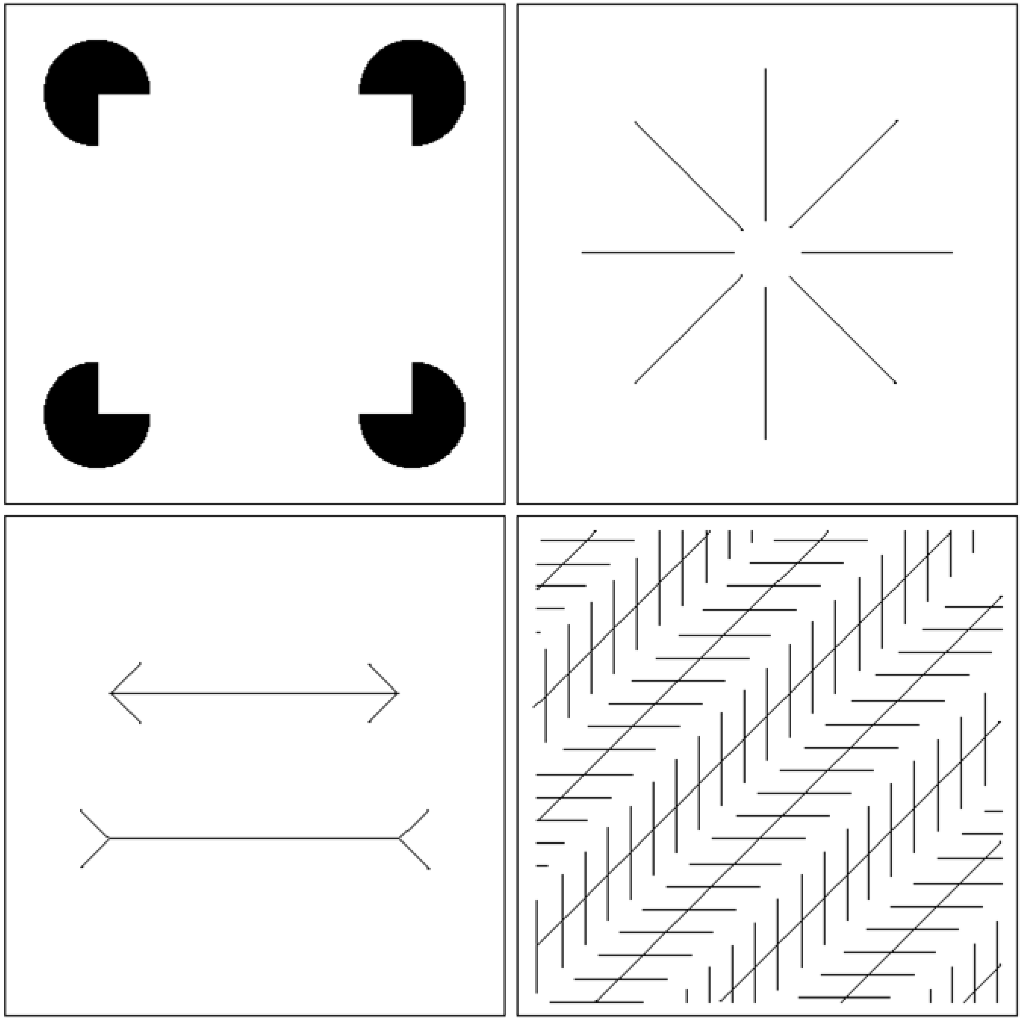

Illustration of Mach band effect
Perceived intensity is not a simple function of the actual intensity

**Course 1**



All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter

**Course 1**



Optical illusions

**Course 1**

| | |
|---|---|
| ALBASTRU | BLUE |
| VERDE | GREEN |
| GALBEN | YELLOW |
| ROSU | RED |
| PORTOCALIU | ORANGE |
| GALBEN | YELLOW |
| VISINIU | BURGUNDY |
| ALB | WHITE |