FPT Education

FPT UNIVERSITY

PART I
IMAGE PROCESSING

INSTRUCTOR: DR. NGUYEN NGOC TRUONG MINH

SCHOOL OF ELECTRICAL ENGINEERING, INTERNATIONAL UNIVERSITY (VNU-HCMC)

Ho Chi Minh City, June 2023

LECTURE I – INTRODUCTION AND OVERVIEW

INSTRUCTOR: DR. NGUYEN NGOC TRUONG MINH

SCHOOL OF ELECTRICAL ENGINEERING, INTERNATIONAL UNIVERSITY (VNU-HCMC)

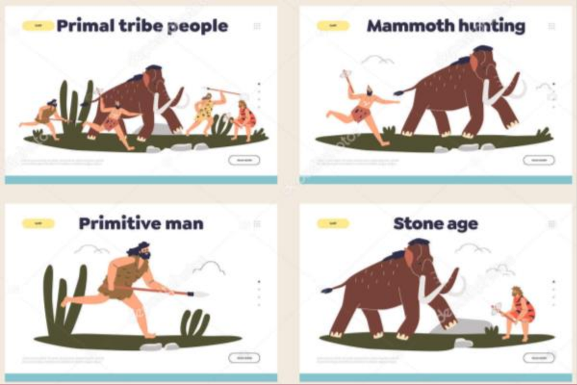
Ho Chi Minh City, June 2023

LECTURE CONTENT

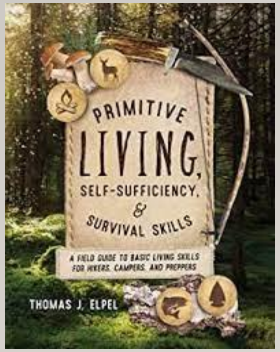
- What is image processing? What are the main applications of image processing?
- What is an image? What is a digital image?
- What are the goals of image processing algorithms?
- What are the most common image processing operations?
- Which hardware and software components are typically needed to build an image processing system?

LECTURE CONTENT

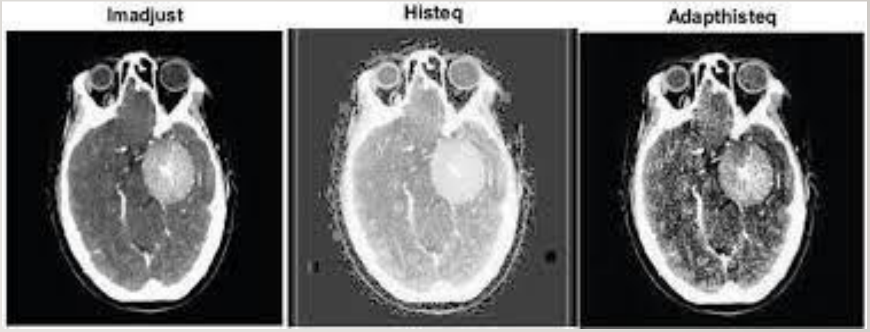
- What is a machine vision system (MVS) and what are its main components?
- Why is it so hard to emulate the performance of the human visual system (HVS) using cameras and computers?
- Lecture Summary – What have we learned?
- Problems



1.1 MOTIVATION

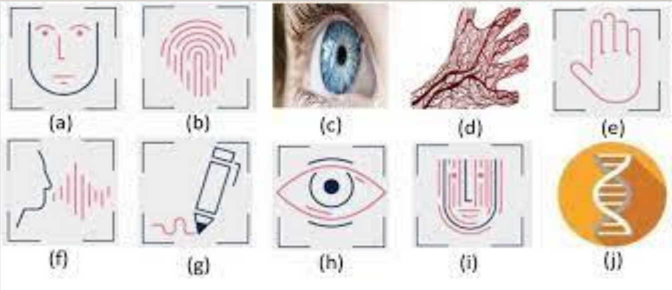


- Humans have historically relied on their vision for tasks ranging from basic instinctive survival skills to detailed and elaborate analysis of works of art. Our ability to guide our actions and engage our cognitive abilities based on visual input is a remarkable trait of the human species, and much of how exactly we do what we do—and seem to do it so well—remains to be discovered.
- The need to extract information from images and interpret their contents has been one of the driving factors in the development of image processing and computer vision during the past decades.



1.1 MOTIVATION

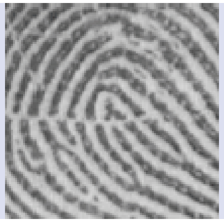
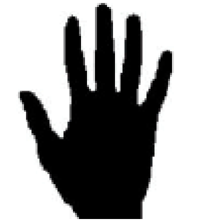
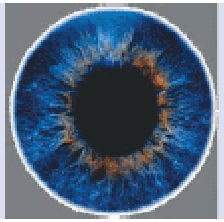
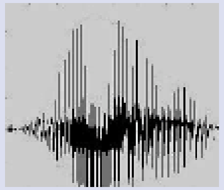
- Image processing applications cover a wide range of human activities, such as the following:
  - *Medical Applications: Diagnostic imaging modalities such as digital radiography, PET (positron emission tomography), CAT (computerized axial tomography), MRI (magnetic resonance imaging), and fMRI (functional magnetic resonance imaging)…*
  - *Industrial Applications: Image processing systems have been successfully used in manufacturing systems for many tasks, such as safety systems, quality control, and control of automated guided vehicles (AGVs)…*



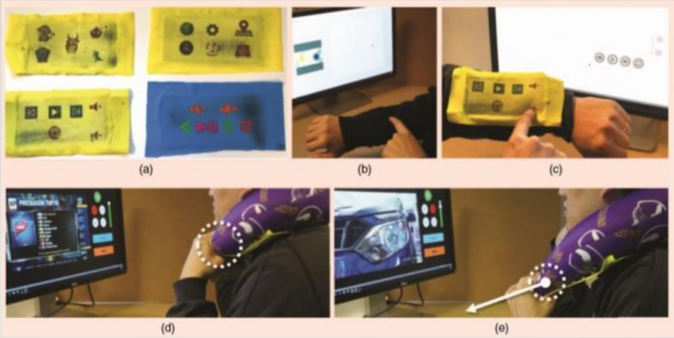
1.1 MOTIVATION

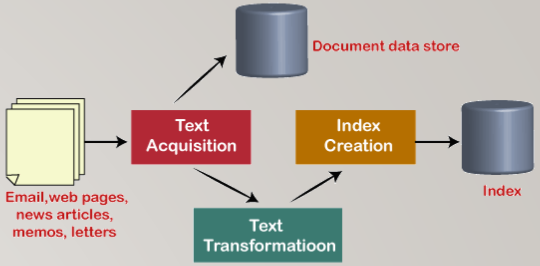
- Image processing applications cover a wide range of human activities, such as the following:
  - *Military Applications: Some of the most challenging and performance-critical scenarios for image processing solutions have been developed for military needs, ranging from detection of soldiers or vehicles to missile guidance and object recognition and reconnaissance tasks using unmanned aerial vehicles (UAVs)…*
  - *Law Enforcement and Security: Surveillance applications have become one of the most intensely researched areas within the video processing community. Biometric techniques (e.g., fingerprint, face, iris, and hand recognition), which have been the subject of image processing research for more than a decade, have recently become commercially available…*

1.1 MOTIVATION

BIOMETRIC	FINGERPRINT	FACE	HAND GEOMETRY	IRIS	VOICE
					
Barriers to universality	Worn ridges; hand or finger impairment	None	Hand impairment	Visual impairment	Speech impairment
Distinctiveness	High	Low	Medium	High	Low
Permanence	High	Medium	Medium	High	Low
Collectibility	Medium	High	High	Medium	Medium
Performance	High	Low	Medium	High	Low
Acceptability	Medium	High	Medium	Low	High
Potential for circumvention	Low	High	Medium	Low	High

1.1 MOTIVATION





The indexing process

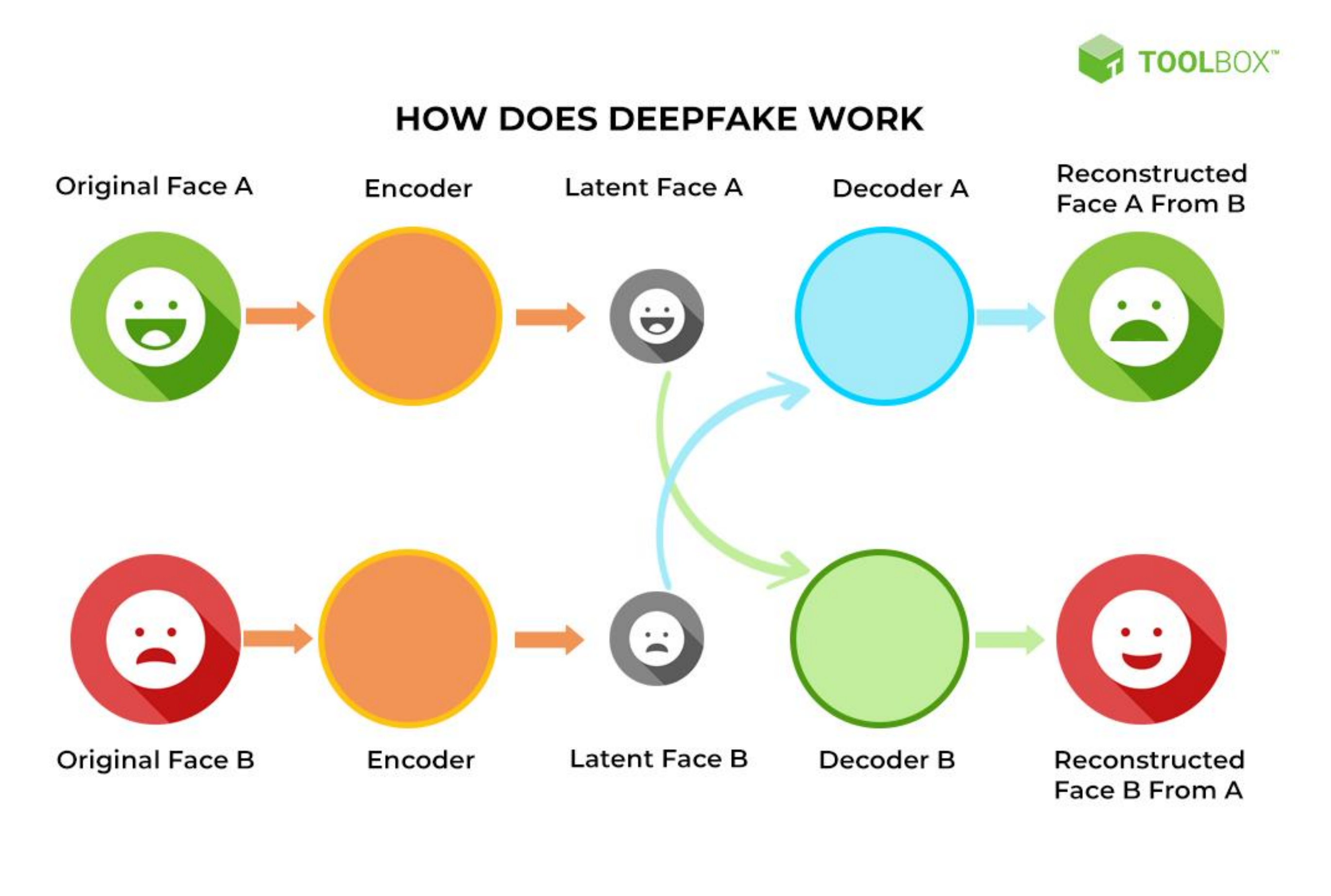
- Image processing applications cover a wide range of human activities, such as the following:
  - *Consumer Electronics: Digital cameras and camcorders, with sophisticated built-in processing capabilities, have rendered film and analog tape technologies obsolete. Software packages to enhance, edit, organize, and publish images and videos have grown in sophistication while keeping a user-friendly interface…*
  - *The Internet, Particularly the World Wide Web: There is a huge amount of visual information available on the Web. Collaborative image and video uploading, sharing, and annotation (tagging) have become increasingly popular. Finding and retrieving images and videos on the Web based on their contents remains an open research challenge.*

1.1 MOTIVATION – FACE ANTI-SPOOFING?

Species	ISO/IEC 30107 Level	FIDO Level
Face image printed on inkjet or laser printer	1	A
Face image printed at photograph laboratory	1	A
Displayed photos on electronic/mobile devices	1	A
Displayed videos on electronic/mobile devices	1	B
Paper masks	2	B
Masks made of specialized materials (ceramic, silicone, and/or theatrical)	3	C
3D printed faces	3	C

1.1 MOTIVATION – DEEPFAKE?

1.1 MOTIVATION – DEEPFAKE?

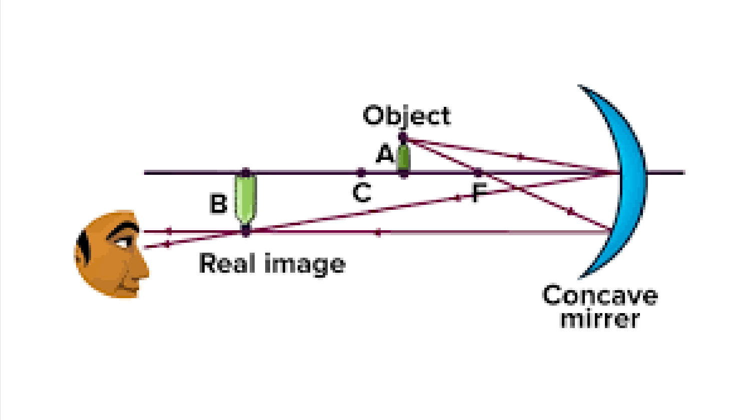


- A deepfake is an artificial intelligence-powered form of media that depicts a person saying something they did not say, appearing in a manner different from authentic visuals, or diverging from reality somehow.
- Its purpose is to fool the media viewer or a technology system. IP can help to explain how deepfakes work, the fraudulent activities they enable, and how to protect against them.

1.2 BASIC CONCEPTS AND TERMINOLOGY

➢ What is an Image?

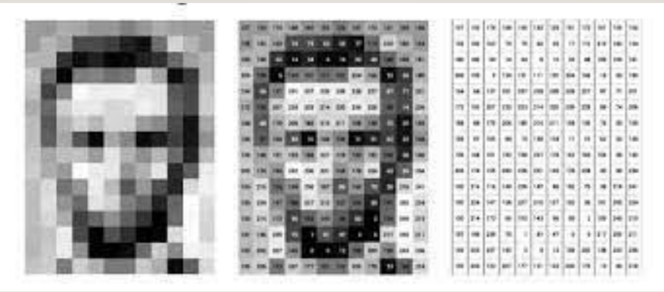
- An image is a visual representation of an object, a person, or a scene produced by an optical device such as a mirror, a lens, or a camera. This representation is two-dimensional (2D), although it corresponds to one of the infinitely many projections of a real-world, three-dimensional (3D) object or scene.



1.2 BASIC CONCEPTS AND TERMINOLOGY

➢ What is a Digital Image?

- A digital image is a representation of a two-dimensional image using a finite number of points, usually referred to as *picture elements*, *pels*, or *pixels*. Each pixel is represented by one or more numerical values: for *monochrome* (grayscale) images, a single value representing the intensity of the pixel (usually in a [0, 255] range) is enough; for color images, three values (e.g., representing the amount of red (R), green (G), and blue (B)) are usually required.



1.2 BASIC CONCEPTS AND TERMINOLOGY

- Monochrome photography is distinguished by variations of a single color, for instance, the green or blue tones. The choice of color in a monochrome shot depends on the personal preferences of a photographer, but only one color is allowed.
- Grayscale photography comprises black and white monochrome colors, with the use of varied gray tones.

1.2 BASIC CONCEPTS AND TERMINOLOGY

Monochrome: YES	Monochrome: YES	Monochrome: YES
Grayscale: YES	Grayscale: NO	Grayscale: NO

1.2 BASIC CONCEPTS AND TERMINOLOGY

1.2 BASIC CONCEPTS AND TERMINOLOGY

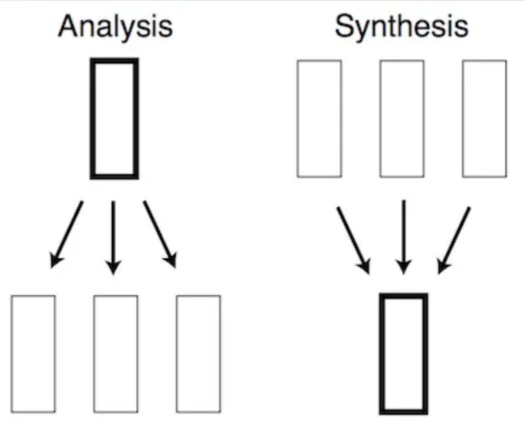
What is Digital Image Processing?

- Digital image processing can be defined as the science of modifying digital images by means of a digital computer. The changes that take place in the images are usually performed automatically and rely on carefully designed algorithms.
- This is in clear contrast with another scenario, such as touching up a photo using an airbrush tool in a photo editing software, in which images are processed manually and the success of the task depends on human ability and dexterity.

1.2 BASIC CONCEPTS AND TERMINOLOGY

➢ What is the Scope of Image Processing?

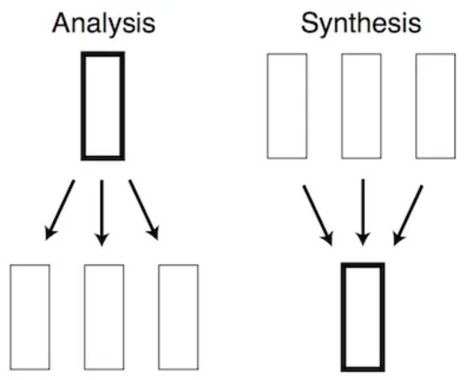
- Whether the output is a modified (i.e., processed) version of the input image, an encoded version of its main attributes, or a nonpictorial description of its contents.
- Low Level: Primitive operations (e.g., noise reduction, contrast enhancement, etc.) where both the input and the output are images.
- Mid Level: Extraction of attributes (e.g., edges, contours, regions, etc.) from images.
- High Level: Analysis and interpretation of the contents of a scene.



1.2 BASIC CONCEPTS AND TERMINOLOGY

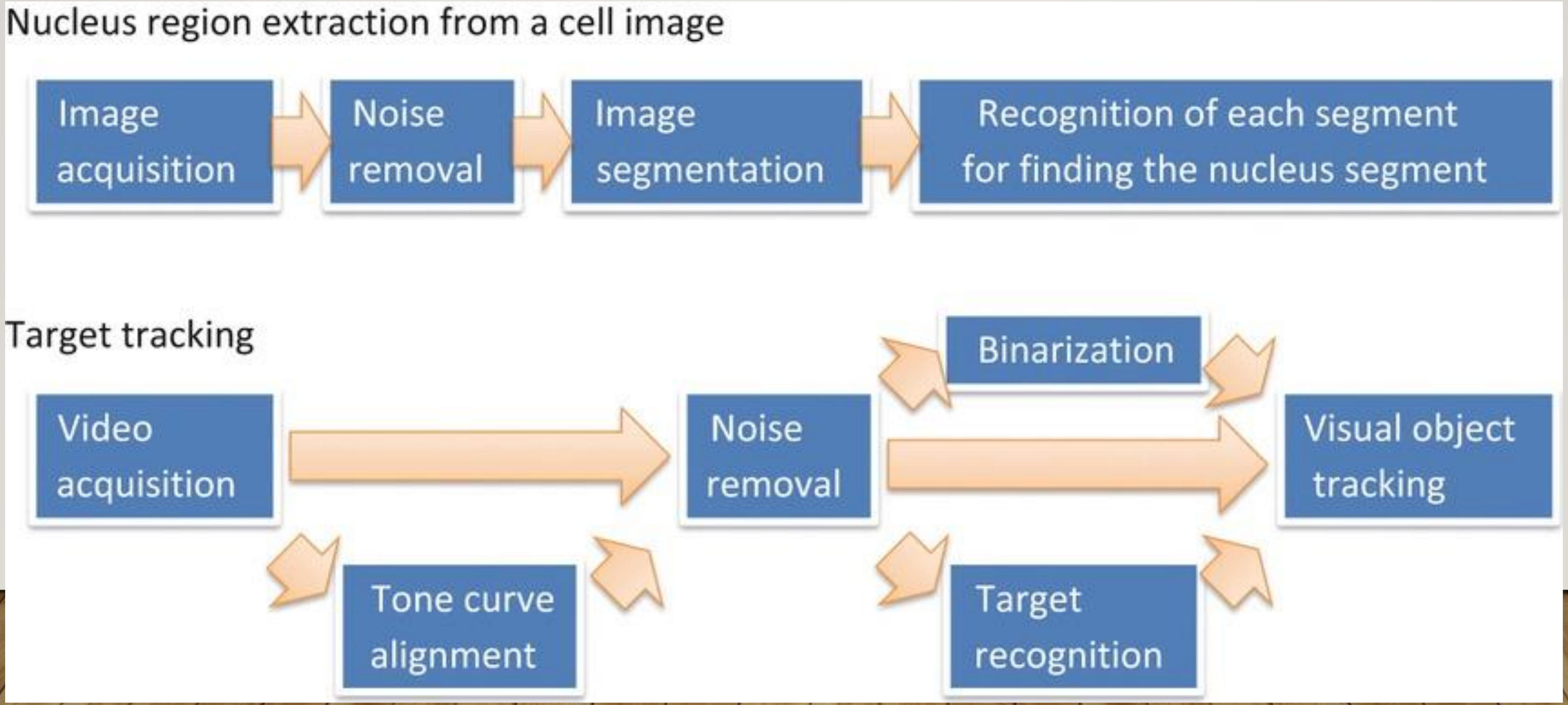
➢ What is the Scope of Image Processing?

- *Image synthesis* (the process by which a 2D or 3D image is rendered from numerical data) vs. *Image analysis* (the process by which textual and numerical data can be extracted from an array of pixel).
- Image processing is a multidisciplinary field, with contributions from different branches of science (particularly mathematics, physics, and computer science) and computer, optical, and electrical engineering.
- Moreover, it overlaps other areas such as pattern recognition, machine learning, artificial intelligence, and human vision research.



1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

- Image processing covers a wide and diverse array of techniques and algorithms ⇨ we provide a preview of the most representative image processing operations that you will learn about in this course.



1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

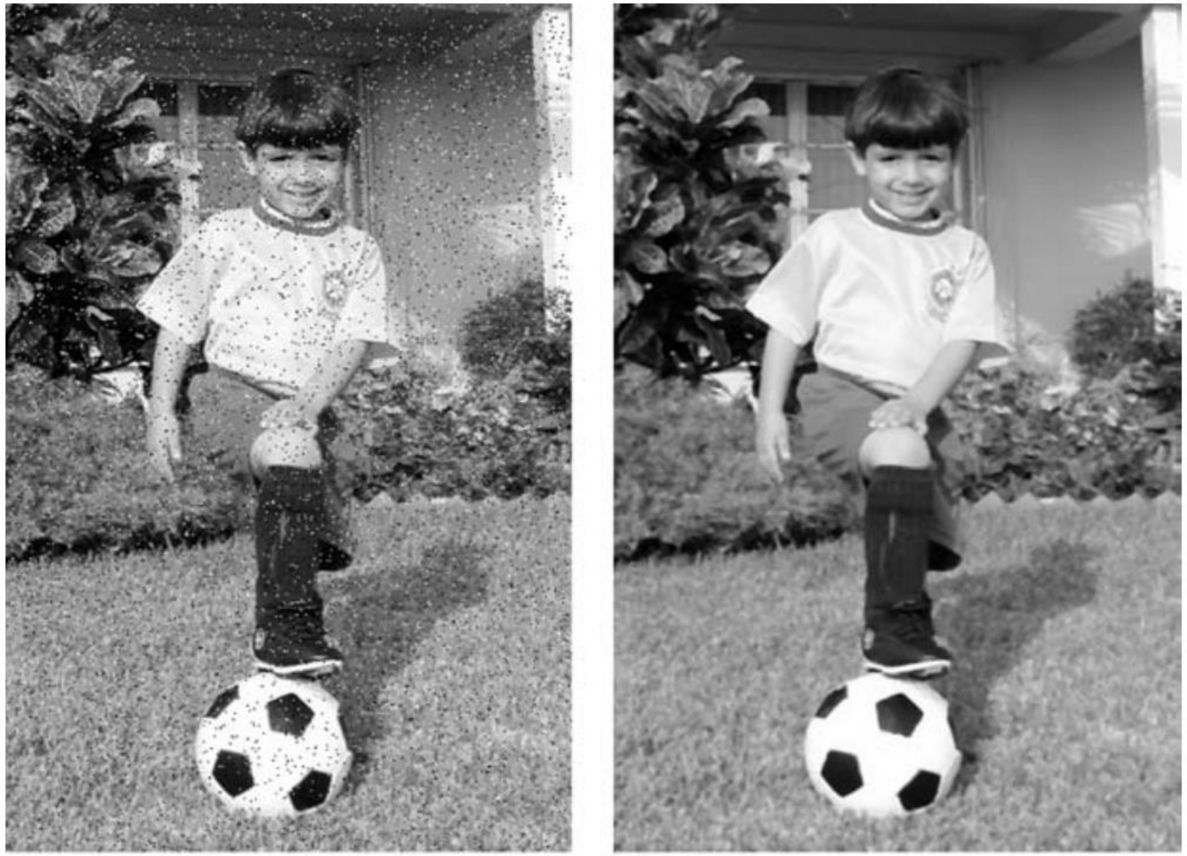
1. Sharpening: A technique by which the edges and fine details of an image are enhanced for human viewing.

(a) (b)

FIGURE 1.1 Image sharpening: (a) original image; (b) after sharpening.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

2. Noise Removal: Image processing filters can be used to reduce the amount of noise in an image before processing it any further. Depending on the type of noise, different noise removal techniques are used.



(a) (b)

FIGURE 1.2 Noise removal: (a) original (noisy) image; (b) after removing noise.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

3. Deblurring: An image may appear blurred for many reasons, ranging from improper focusing of the lens to an insufficient shutter speed for a fast-moving object.

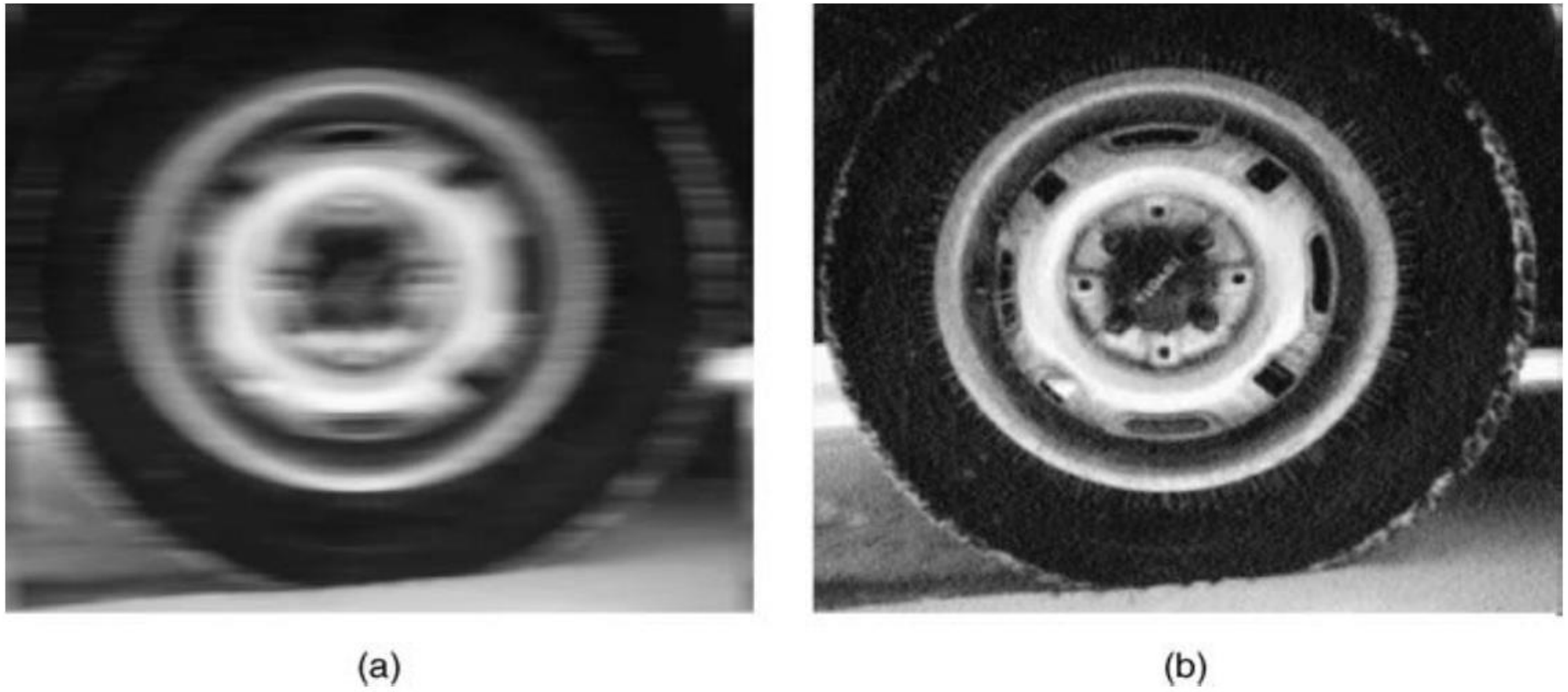


FIGURE 1.3 Deblurring: (a) original (blurry) image; (b) after removing the (motion) blur. Original image: courtesy of MathWorks.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

4. Edge Extraction:

Extracting edges from an image is a fundamental preprocessing step used to separate objects from one another before identifying their contents.

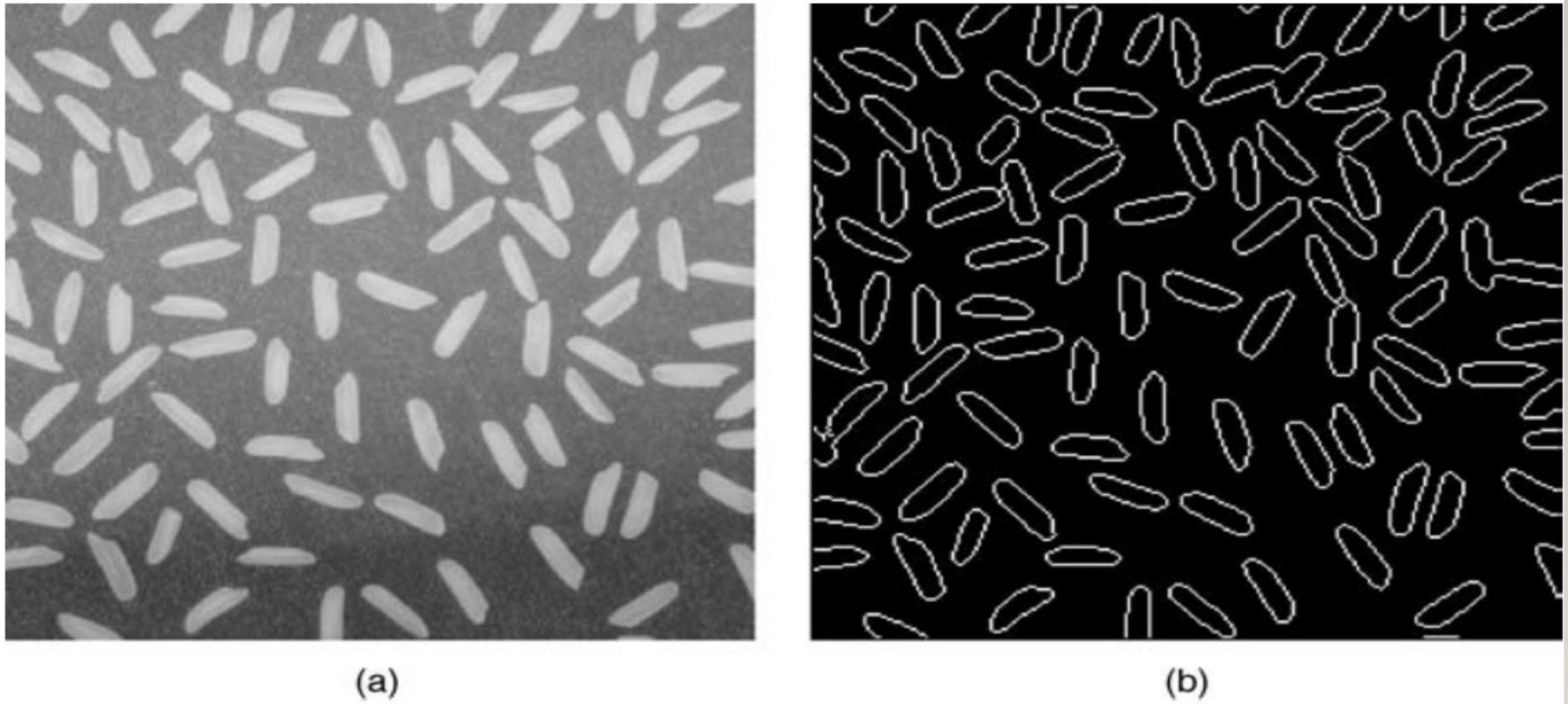


FIGURE 1.4 Edge extraction: (a) original image; (b) after extracting its most relevant edges. Original image: courtesy of MathWorks.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

5. Binarization: In many image analysis applications, it is often necessary to reduce the number of gray levels in a monochrome image to simplify and speed up its interpretation. Reducing a grayscale image to only two levels of gray (black and white) is usually referred to as *binarization*.

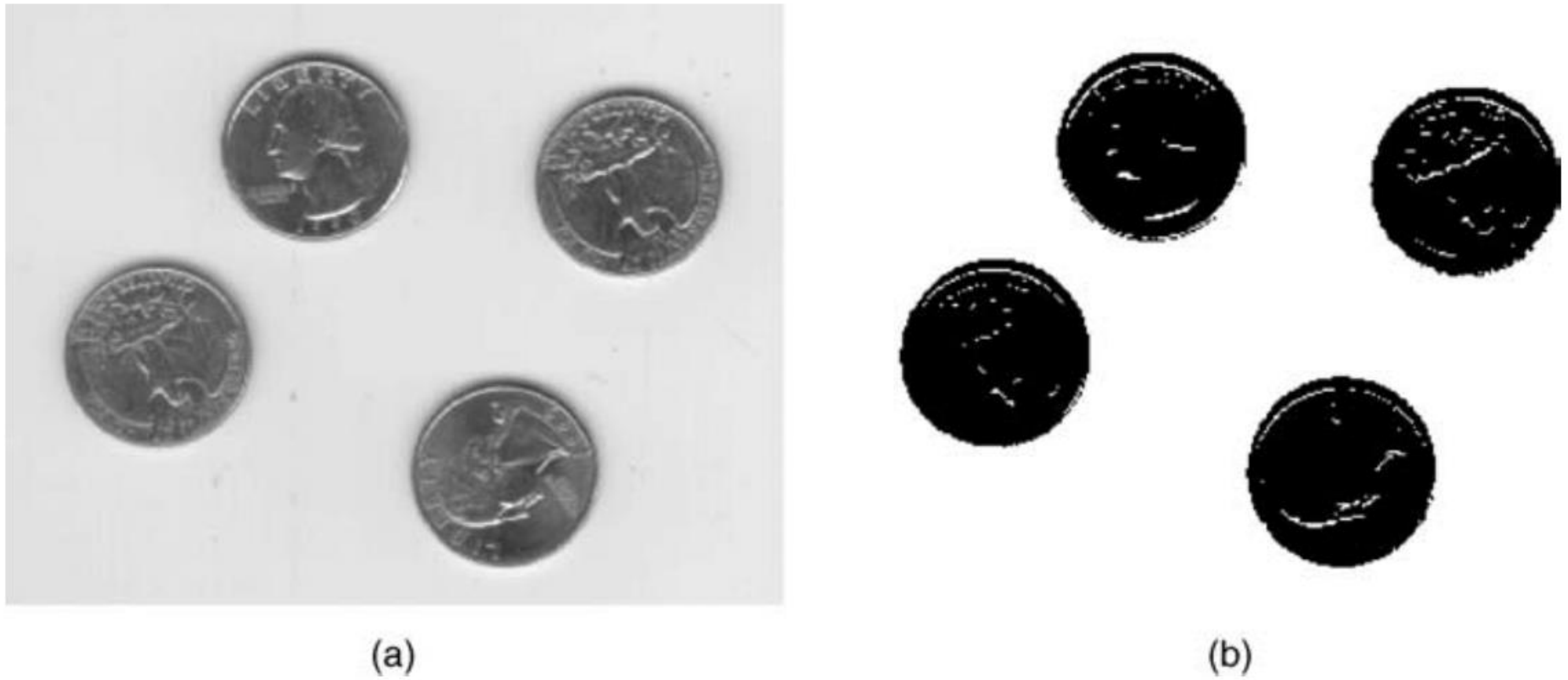


FIGURE 1.5 Binarization: (a) original grayscale image; (b) after conversion to a black-and-white version. Original image: courtesy of MathWorks.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

6. Blurring: It is sometimes necessary to blur an image in order to minimize the importance of texture and fine detail in a scene, for instance, in cases where objects can be better recognized by their shape. Blurring techniques in spatial and frequency domain.

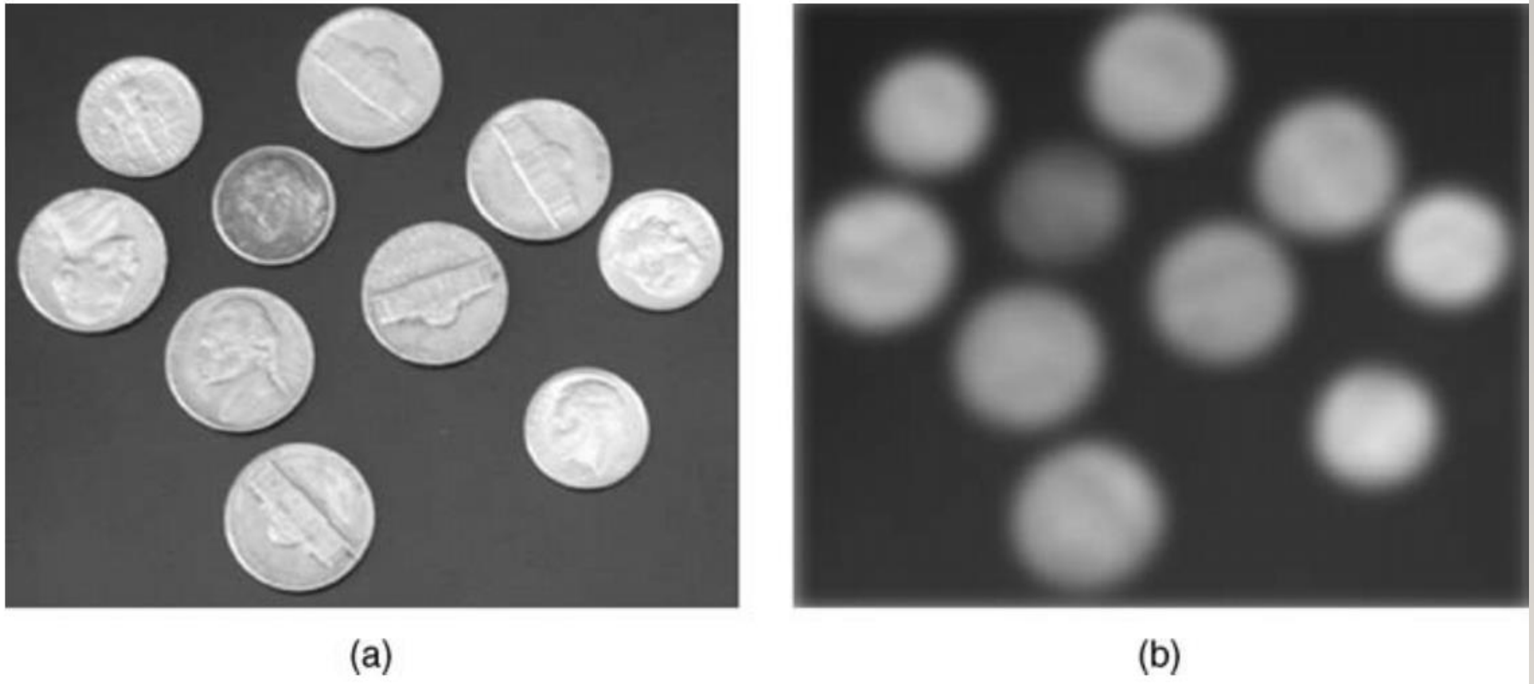


FIGURE 1.6 Blurring: (a) original image; (b) after blurring to remove unnecessary details. Original image: courtesy of MathWorks.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

7. Contrast Enhancement: In order to improve an image for human viewing as well as make other image processing tasks (e.g., edge extraction) easier, it is often necessary to enhance the contrast of an image. Contrast enhancement techniques using transformation functions and histogram processing.

(a) (b)

FIGURE 1.7 Contrast enhancement: (a) original image; (b) after histogram equalization to improve contrast.

1.3 EXAMPLES OF TYPICAL IMAGE PROCESSING OPERATIONS

8. Object Segmentation and Labeling:

The task of segmenting and labeling objects within a scene is a prerequisite for most object recognition and classification systems. Once the relevant objects have been segmented and labeled, their relevant features can be extracted and used to classify, compare, cluster, or recognize the objects in question. Segmentation and labeling of connected components from an image.

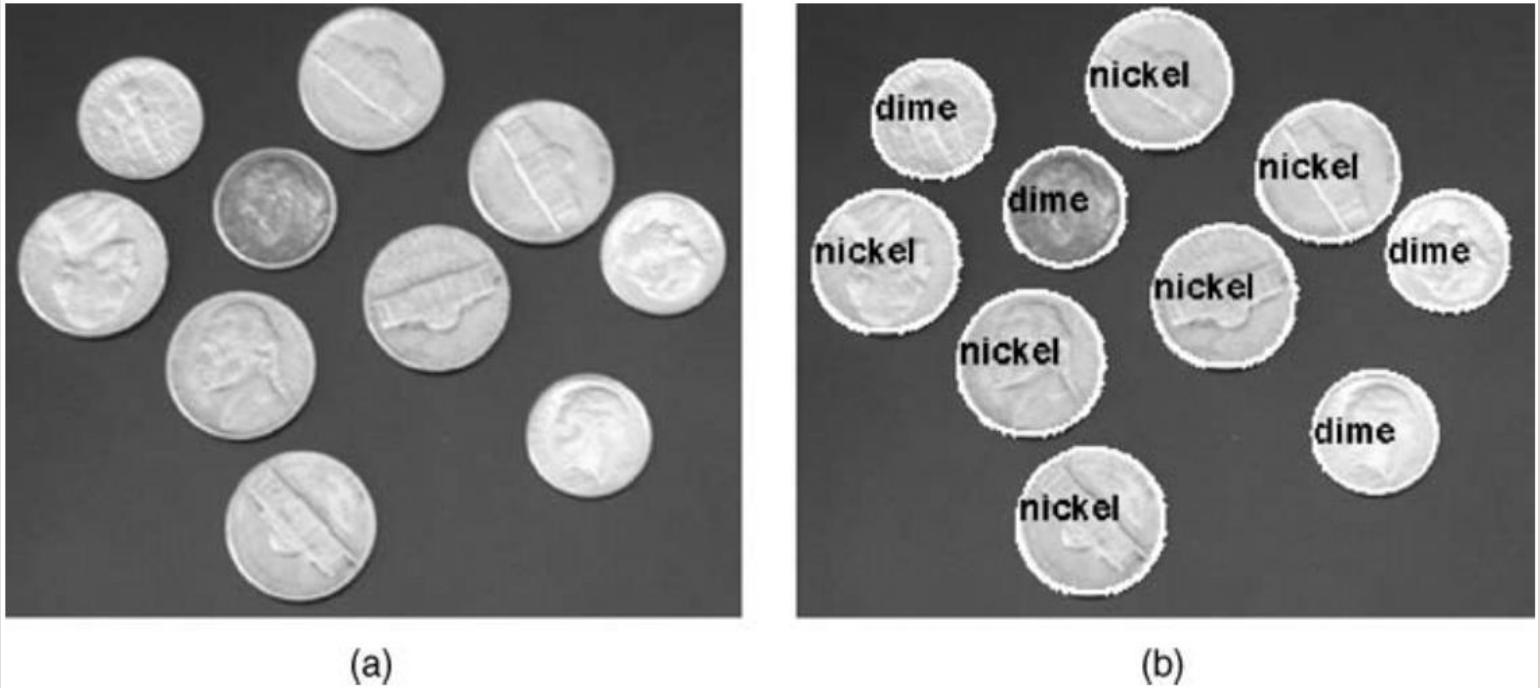


FIGURE 1.8 Object segmentation and labeling: (a) original image; (b) after segmenting and labeling individual objects. Original image: courtesy of MathWorks.

1.4 COMPONENTS OF A DIGITAL IMAGE PROCESSING SYSTEM

- The system is built around a computer in which most image processing tasks are carried out, but also includes hardware and software for image acquisition, storage, and display.
- In fact, even contemporary digital still cameras can be modeled according to that diagram: the CCD sensor corresponds to the Acquisition block, flash memory is used for storage, a small LCD monitor for display, and the digital signal processor (DSP) chip becomes the 'Computer', where certain image processing operations (e.g., conversion from RAW format to JPEG) take place.

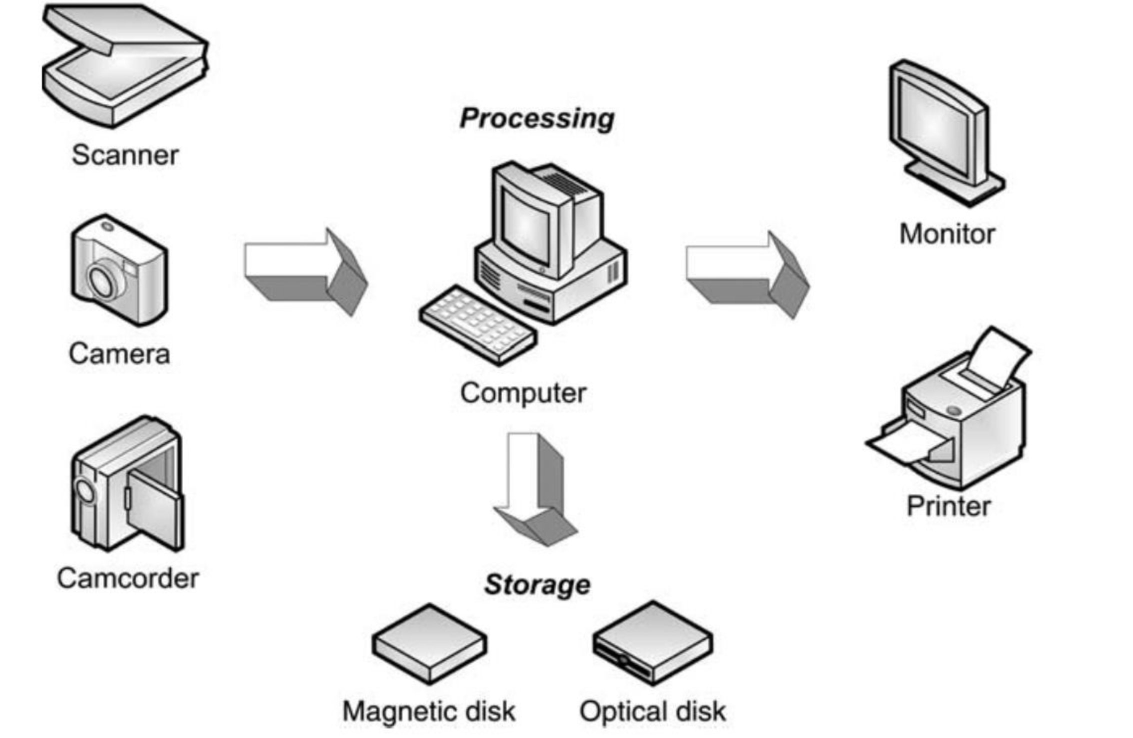
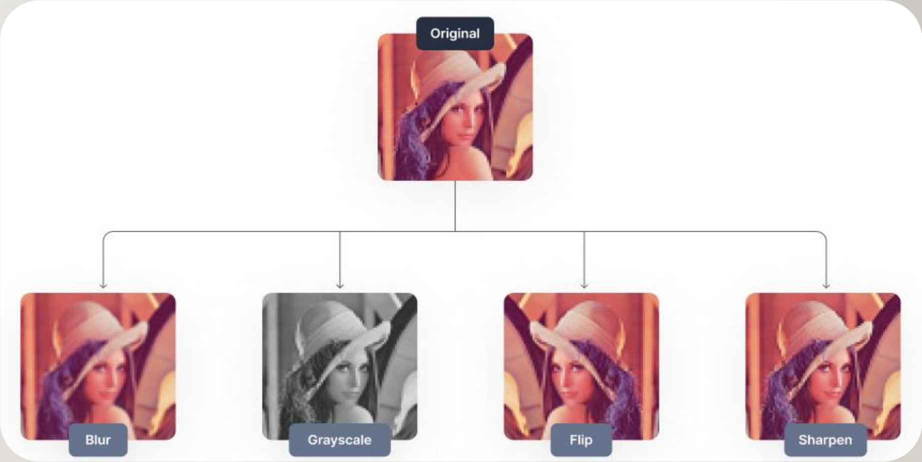


FIGURE 1.9 Components of a digital image processing system. Adapted and redrawn from [Umb05].

1.4 COMPONENTS OF A DIGITAL IMAGE PROCESSING SYSTEM - HARDWARE

- The hardware components of a digital image processing system typically include the following:
- *Acquisition Devices*: Responsible for capturing and digitizing images or video sequences. Examples of general-purpose acquisition devices include scanners, cameras, and camcorders.
- *Processing Equipment*: The main computer itself, in whatever size, shape, or configuration. Responsible for running software that allows the processing and analysis of acquired images.
- *Display and Hardcopy Devices*: Responsible for showing the image contents for human viewing. Examples include color monitors and printers.
- *Storage Devices*: Magnetic or optical disks responsible for long-term storage of the images.

1.4 COMPONENTS OF A DIGITAL IMAGE PROCESSING SYSTEM – HARDWARE

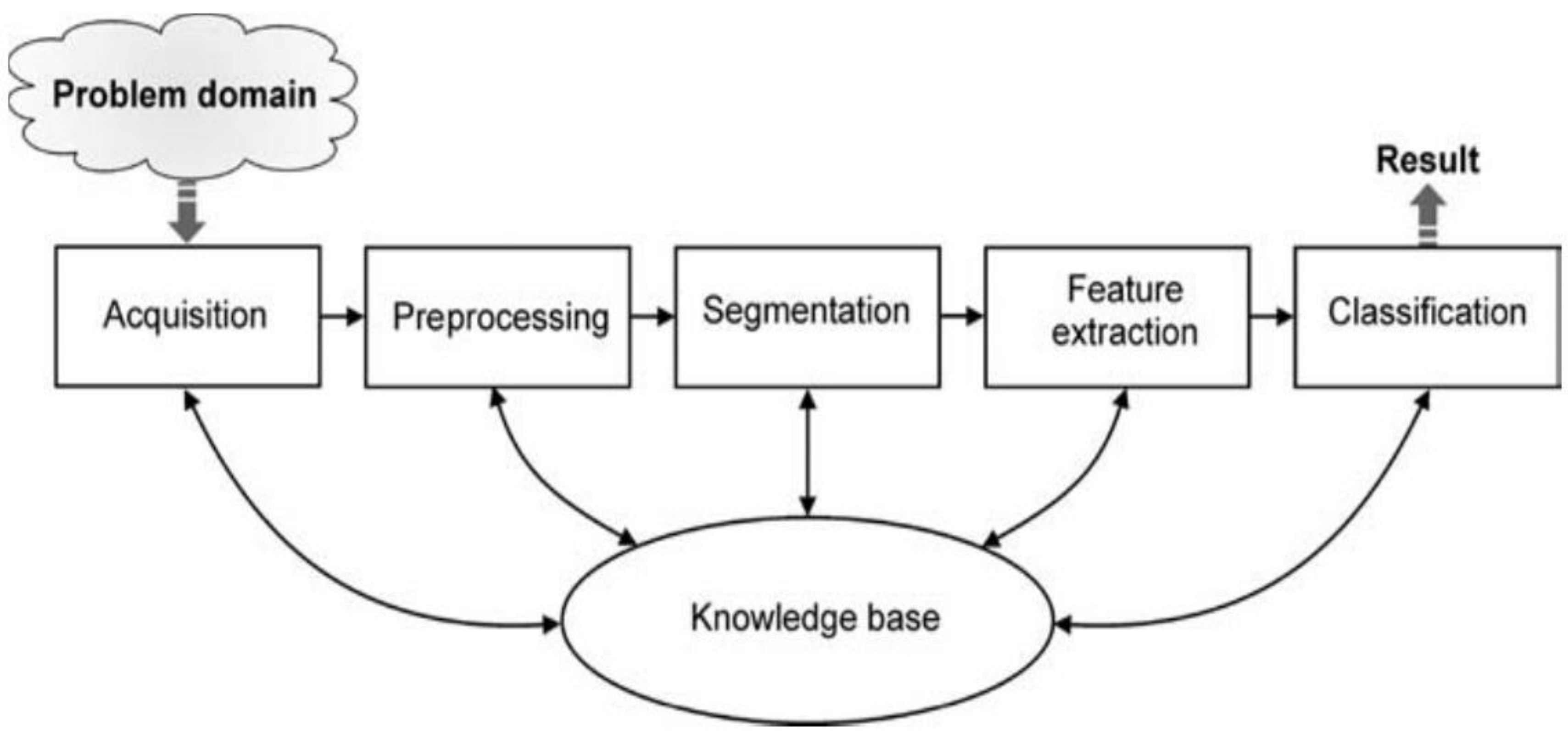


1.4 COMPONENTS OF A DIGITAL IMAGE PROCESSING SYSTEM – SOFTWARE

- The software portion of a digital image processing system usually consists of modules that perform specialized tasks. The development and fine-tuning of software for image processing solutions is iterative in nature.
- Consequently, image processing researchers and practitioners rely on programming languages and development environments that support modular, agile, and iterative software development.
- MATLAB has become very popular with engineers, scientists, and researchers in both industry and academia, due to many factors, such as the availability of toolboxes containing specialized functions for many application areas, ranging from data acquisition to image processing.

1.5 MACHINE VISION SYSTEMS

- The main components of a machine vision system using a practical example application: recognizing license plates at a high-way toll booth.
- Image processing is not a one-step process: most solutions follow a sequential processing scheme.



1.5 MACHINE VISION SYSTEMS

- ***The problem domain***, in this case, is the automatic recognition of license plates. The goal is to be able to extract the alphanumeric contents of the license plate of a vehicle passing through the toll booth in an automated and unsupervised way, that is, without need for human intervention.
- Additional requirements could include *24/7 operation* (under artificial lighting), *all-weather operation*, *minimal acceptable success rate*, and *minimum and maximum vehicle speed*.

1.5 MACHINE VISION SYSTEMS

- *The acquisition block* is in charge of *acquiring one or more images* containing a front or rear view of the vehicle that includes its license plate.

- This can be implemented using a CCD camera and controlling the lighting conditions so as to ensure that the image will be suitable for further processing. *The output of this block is a digital image that contains a (partial) view of the vehicle.*

1.5 MACHINE VISION SYSTEMS

- The goal of ***the preprocessing stage*** is to improve the quality of the acquired image.
- Possible algorithms to be employed during this stage include *contrast improvement*, *brightness correction*, and *noise removal*.

1.5 MACHINE VISION SYSTEMS

- *The segmentation block* is responsible for partitioning an image into its main components: relevant foreground objects and background. It produces at its output a number of labeled regions or "sub-images".
- It is possible that in this particular case segmentation will be performed at two levels: (1) extracting the license plate from the rest of the original image; and (2) segmenting characters within the plate area.
- Automatic image segmentation is *one of the most challenging tasks in a machine vision system*.

1.5 MACHINE VISION SYSTEMS

- *The feature extraction block* (also known as representation and description) consists of algorithms responsible for encoding the image contents in a concise and descriptive way.
- Typical features include measures of color (or intensity) distribution, texture, and shape of the most relevant (previously segmented) objects within the image.
- These features are usually grouped into a feature vector that can then be used as *a numerical indicator of the image (object) contents* for the subsequent stage, where such contents will be recognized (*classified*).

1.5 MACHINE VISION SYSTEMS

- Once the most relevant features of the image (or its relevant objects, in this case individual characters) have been extracted and encoded into a feature vector, the next step is to use this K-dimensional numerical representation as an input to ***the pattern classification*** (also known as recognition and interpretation) stage.
- At this point, image processing meets classical pattern recognition and benefits from many of its tried-and-true techniques, such as *minimum distance classifiers*, *probabilistic classifiers*, *neural networks*, and *many more*.

1.5 MACHINE VISION SYSTEMS

- All modules are connected to a large block called ***knowledge base***.
- These connections are meant to indicate that the successful solution to the license plate recognition problem will depend on how much knowledge about the problem domain has been encoded and stored in the MVS.
- The role of such knowledge base in the last stages is quite evident knowledge that the first character must be a digit may help disambiguate between a “0” and an “O” in the pattern classification stage.

1.5 MACHINE VISION SYSTEMS

- The human visual system (HVS) and a machine vision system (MVS) have different strengths and limitations.
- Three of the biggest challenges stand out:
  - The HVS can rely on a very large database of images and associated concepts that have been captured, processed, and recorded during a lifetime.
  - The very high speed at which the HVS makes decisions based on visual input. Although several image processing and machine vision tasks can be implemented at increasingly higher speeds.
  - The remarkable ability of the HVS to work under a wide range of conditions, from deficient lighting to less-than-ideal perspectives for viewing a 3D object.

WHAT HAVE WE LEARNED?

- *Digital image processing* is the science of modifying digital images using a digital computer.
- Digital image processing is closely related to other areas such as *computer vision* and *pattern recognition*.
- *Digital image processing algorithms*, *techniques*, and *applications* usually take an image as input and produce one of the following outputs: a modified (i.e., processed) image, an encoded version of the main attributes present in the input image, or a nonpictorial description of the input image's contents.

WHAT HAVE WE LEARNED?

- *Digital image processing* has found applications in almost every area of modern life, from medical imaging devices to quality control in manufacturing systems, and from consumer electronics to law enforcement and security.
- An image is *a visual representation of an object*, *a person*, or *a scene* produced by an optical device such as a mirror, a lens, or a camera. This representation is two dimensional, although it corresponds to one of the infinitely many projections of a real-world, three-dimensional object or scene.

WHAT HAVE WE LEARNED?

- *A digital image* is a representation of a two-dimensional image using a finite number of pixels, each of which indicates the gray level or color contents of the image at that point.
- *Image manipulation techniques* consist of manually modifying the contents of an image using preexisting tools (e.g., airbrush).
- Representative image processing operations include image sharpening, noise removal, edge extraction, contrast enhancement, and object segmentation and labeling.

WHAT HAVE WE LEARNED?

- *A digital image processing system* is usually built around a general-purpose computer equipped with hardware for image and video acquisition, storage, and display. The software portion of the system usually consists of modules that perform specialized tasks. In this book, we shall use MATLAB (and its Image Processing Toolbox) as the software of choice.
- *A machine vision system* is a combination of hardware and software designed to solve problems involving the analysis of visual scenes using intelligent algorithms. Its main components are acquisition, preprocessing, segmentation, feature extraction, and classification.
- It is extremely difficult to emulate the performance of *the human visual system* - in terms of processing speed, previously acquired knowledge, and the ability to resolve visual scenes under a wide range of conditions using machine vision systems.

PROBLEMS

Problem 1.1. Use the block diagram from Figure 1.10 as a starting point to design a machine vision system to read the label of the main integrated circuit (IC) on a printed circuit board (PCB) (see Figure 1.11 for an example). Explain what each block will do, their input and output, what are the most challenging requirements, and how they will be met by the designed solution.

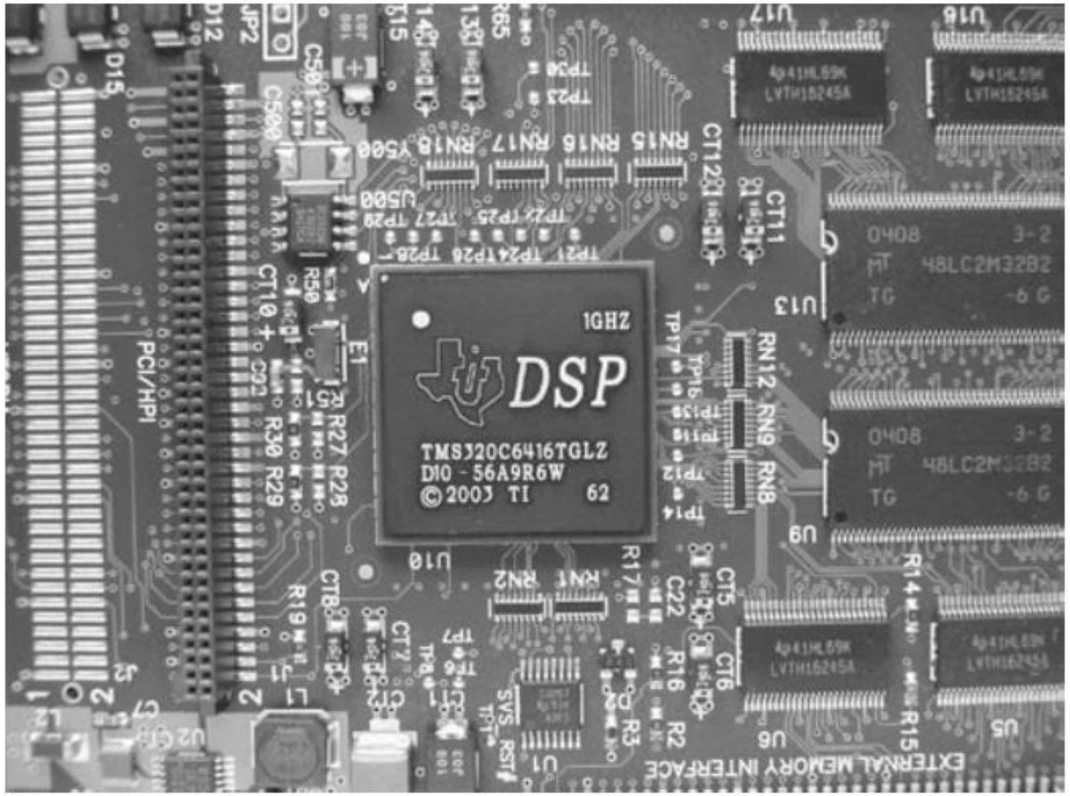


FIGURE 1.11 Test image for the design of a machine vision system to read the label of the main integrated circuit on a printed circuit board.

PROBLEMS

Problem 1.2. In our discussion on machine vision systems, we indicated that the following are the three biggest difficulties in emulating the human visual system: its huge database (images and concepts captured, processed, and recorded during a lifetime), its high speed for processing visual data and making decisions upon them, and the ability to perform under a wide range of work conditions. Explain each of these challenges in your own words, and comment on which ones are more likely to be minimized, thanks to advances in image processing hardware and software.

Problem 1.3. Who do you think would perform better at the following tasks: man (HVS) or computer (MVS)? Please explain why.

PROBLEMS

(a) Determining which line is the shortest in Figure 1.12a.

(b) Determining which circle is the smallest in Figure 1.12b.

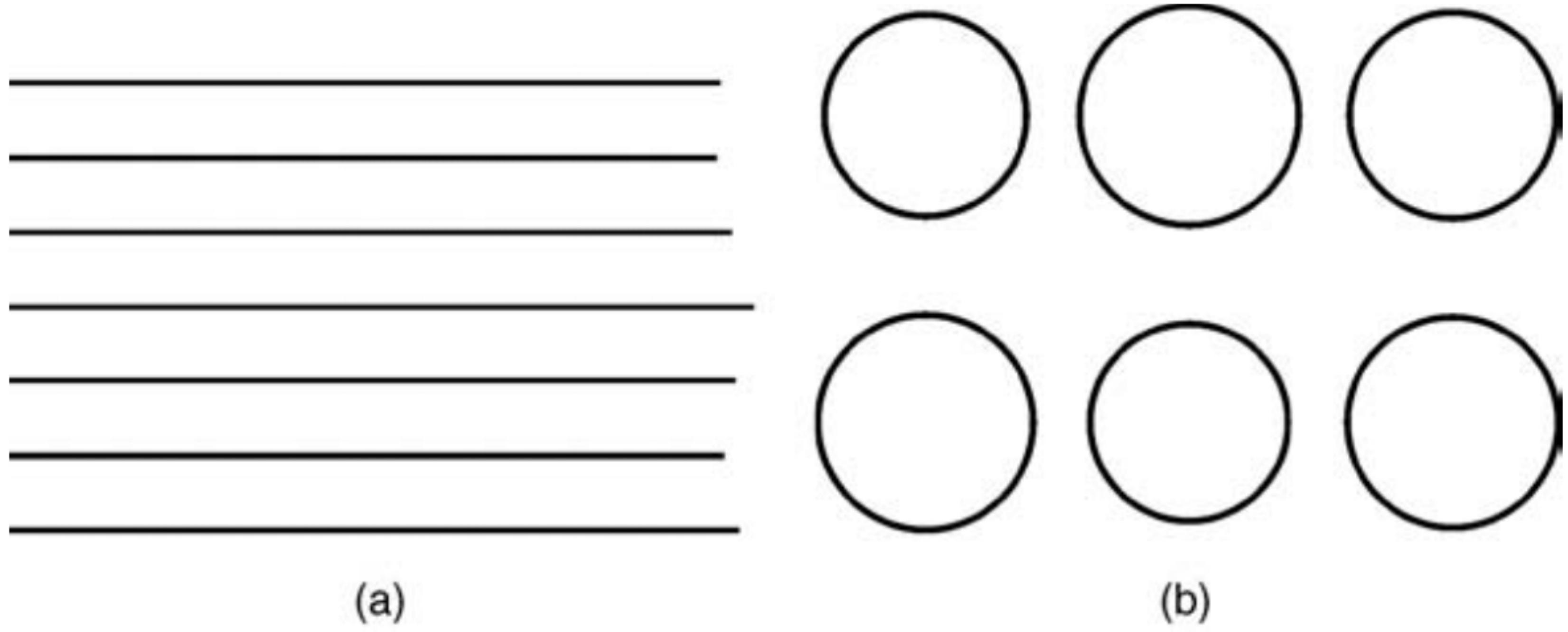


FIGURE 1.12 (a) Test image for distance estimation: parallel lines with up to 5% difference in length. (b) Test image for area estimation: circles with up to 10% difference in radius. Both images are adapted and redrawn from [Jah05].

PROBLEMS

(a) Segmenting the image containing the letter “F” from the background in Figure 1.13a.

(b) Segmenting the white triangle (this triangle—known as “Kanizsa’s triangle”—is a well-known optical illusion) in Figure 1.13b.

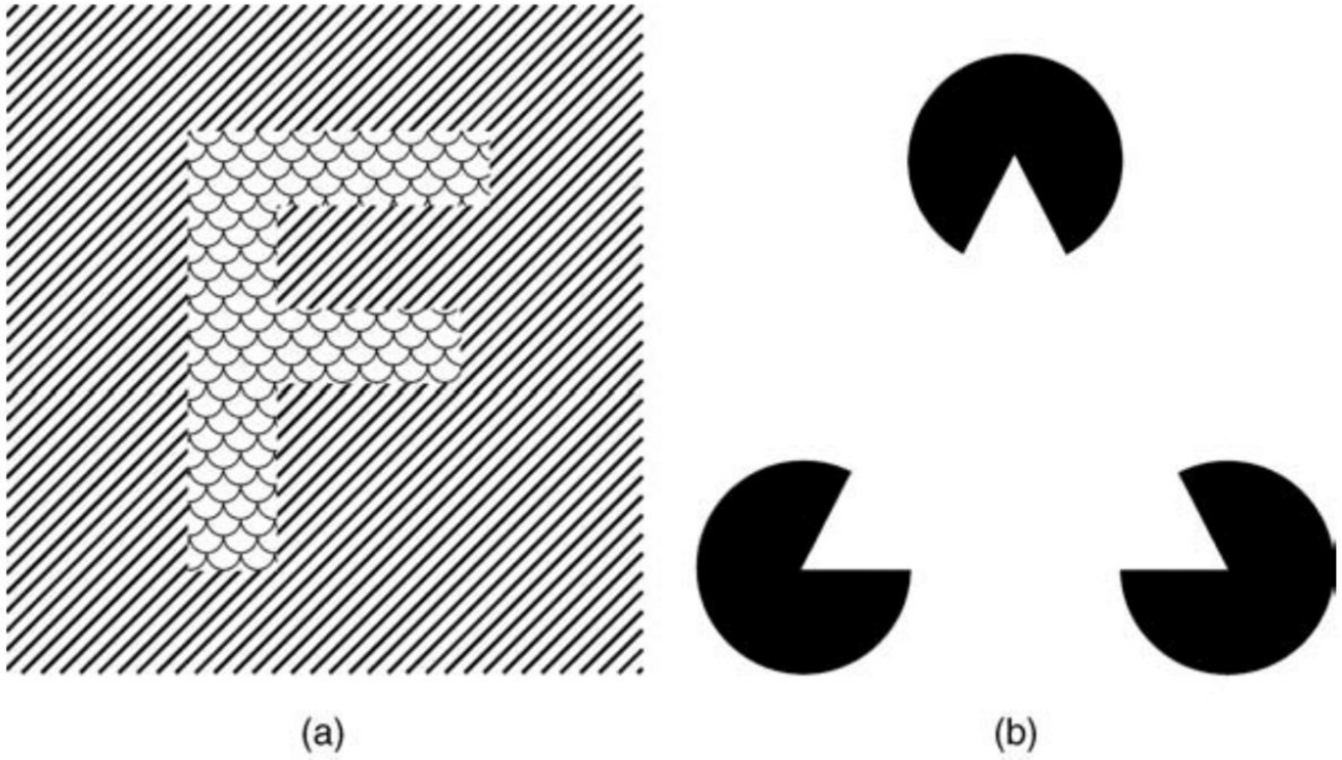


FIGURE 1.13 (a) Test image for texture-based object segmentation. (b) Test image for object segmentation based on “interpolation” of object boundaries. Both images are adapted and redrawn from [Jah05].

END OF LECTURE 1

LECTURE 2 – IMAGE PROCESSING BASICS