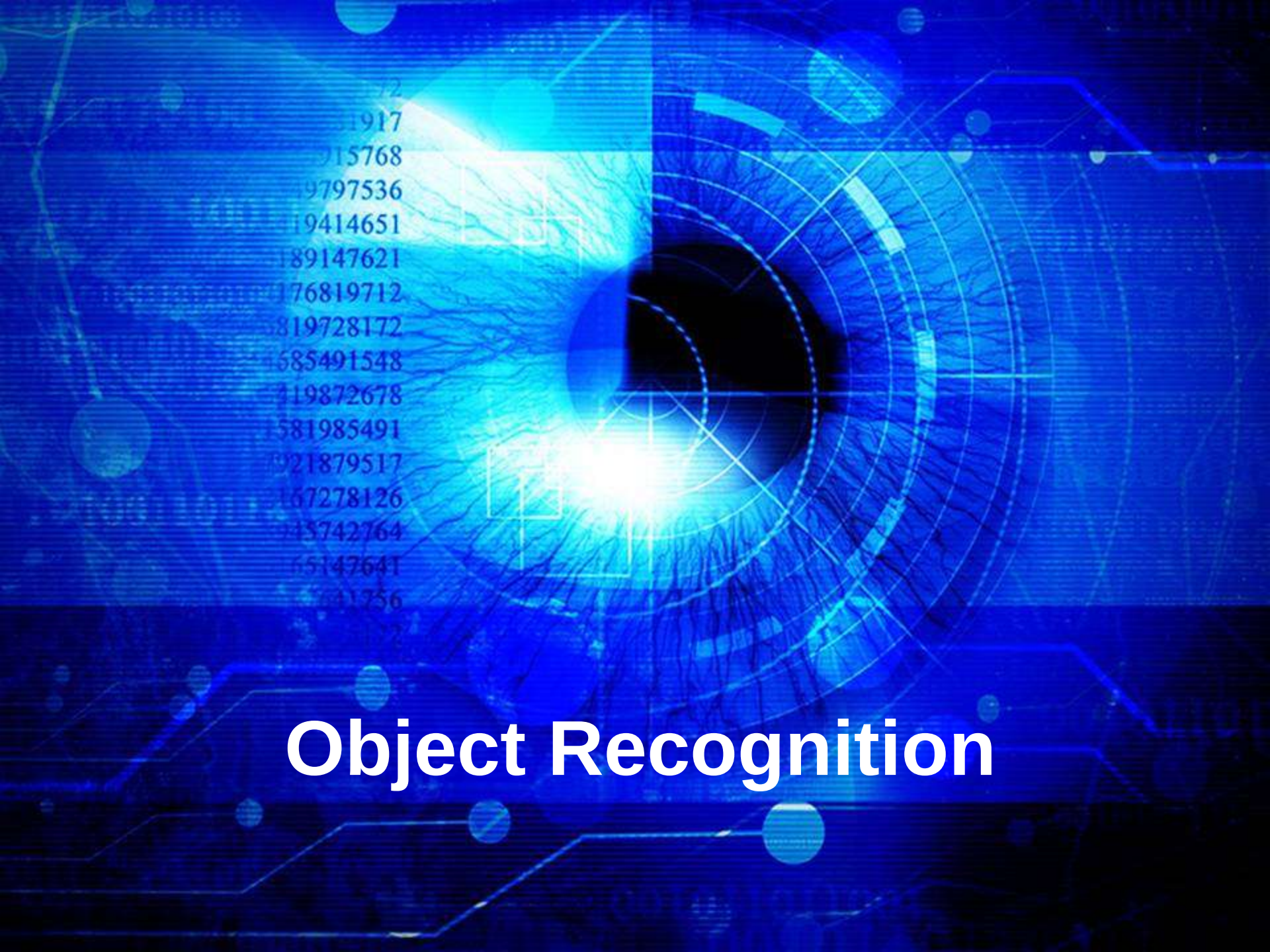

Object Recognition

# Outline

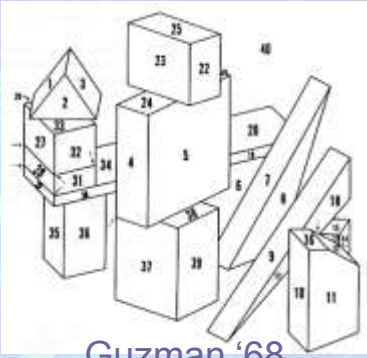
- Introduction (Computer Vision)
  - History
  - Human Vision Vs. Computer Vision
  - Main Goal of Computer Vision
  - Significance of Computer Vision
  - Connections to other Disciplines
  - Key Stages in Digital Image Processing
- Object Recognition
  - What is Object Recognition?
  - What is Pattern Recognition?
    - Approaches
    - Applications
    - Main Components
    - Gender Example

# Outline

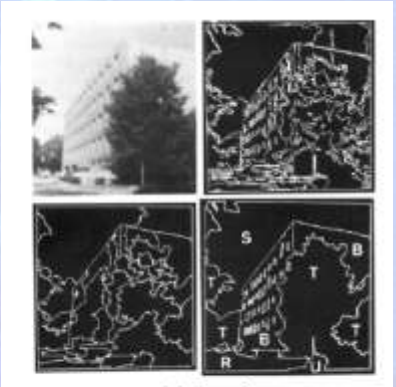
- Fingerprint Recognition
  - Definition
  - Fingerprint Matching Using Ridge-End and Burification
  - Fingerprint Image
    - Binarization
    - Thinning
  - Minutiae Extraction
- Car Number Plate Recognition
  - What is an ALPR System?
  - ALPR Procedure
    - Characters Recognition
    - Characters Segmentation
    - Normalization of Characters
- New Innovations in Object Recognition
- References

# Brief History of Computer Vision

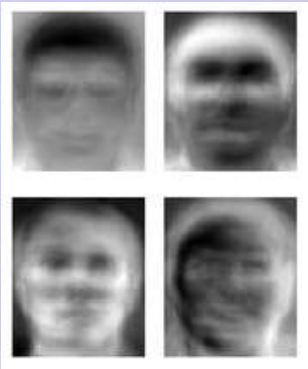
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2030's: robot uprising?



Guzman '68



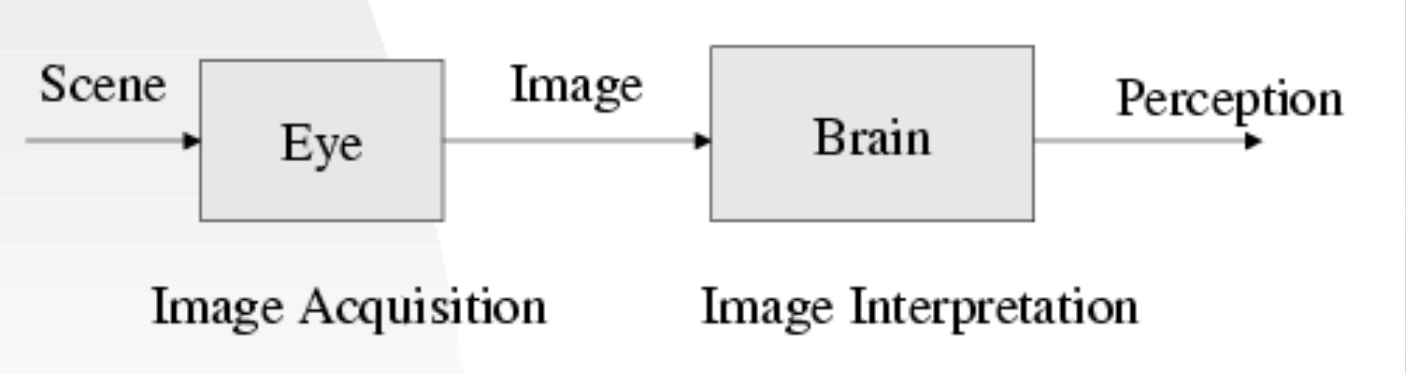
Ohta Kanade '78

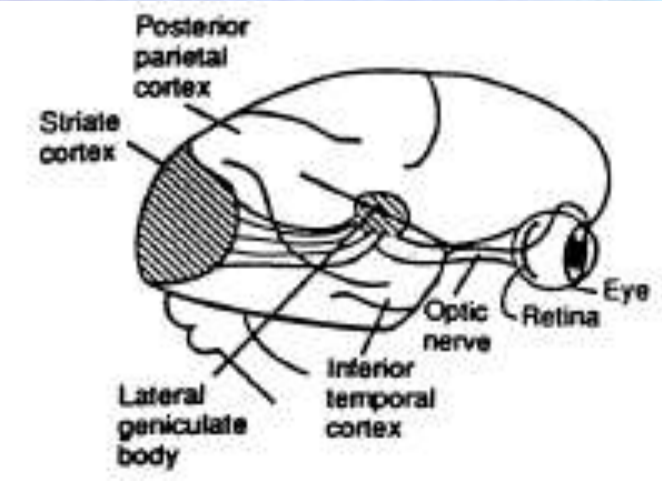


Turk and Pentland '91

# Human Vision

- Vision is the process of discovering what is present in the world and where it is by looking

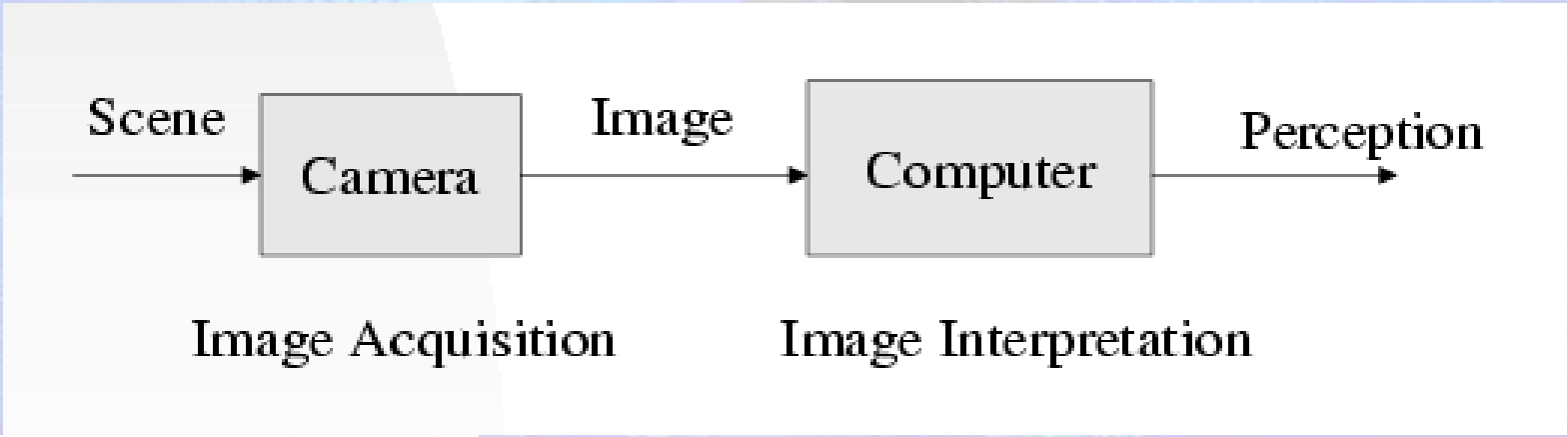


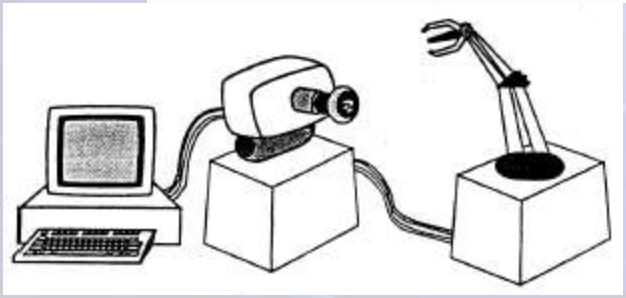




- Computational algo implemented in this massive network of neurons; they obtain their inputs from retina, & produce as output an "understanding" of the scene in view
  - But what does it mean to "understand" the scene? What algos & data representation are used by brain?

# Computer Vision

- Computer Vision is the study of analysis of pictures and videos in order to achieve results similar to those as by humans



- Analogously, given a set of TV camera
  - What computer architectures, data structures & algorithms should use to create a machine that can “see” as we do?

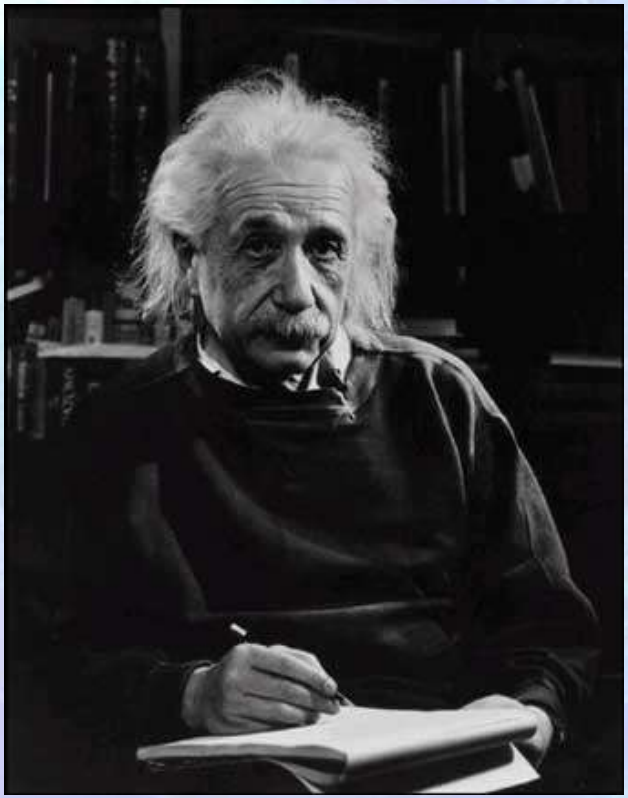
# Human Vision VS Computer Vision


| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 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 we see**

**What a computer sees**

# Main Goal of Computer Vision



**Every picture tells a story!!**

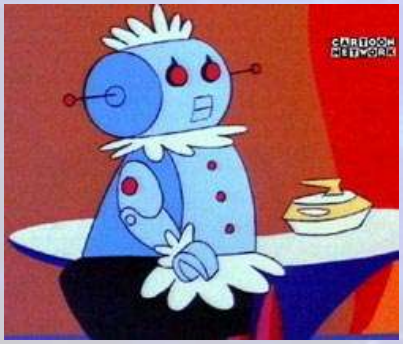
*** write computer programs that can interpret images ***

# Significance of Computer Vision

Safety

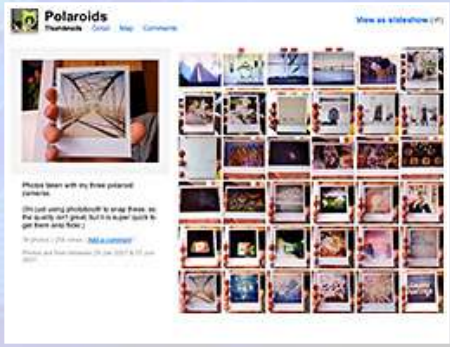

Health

Security



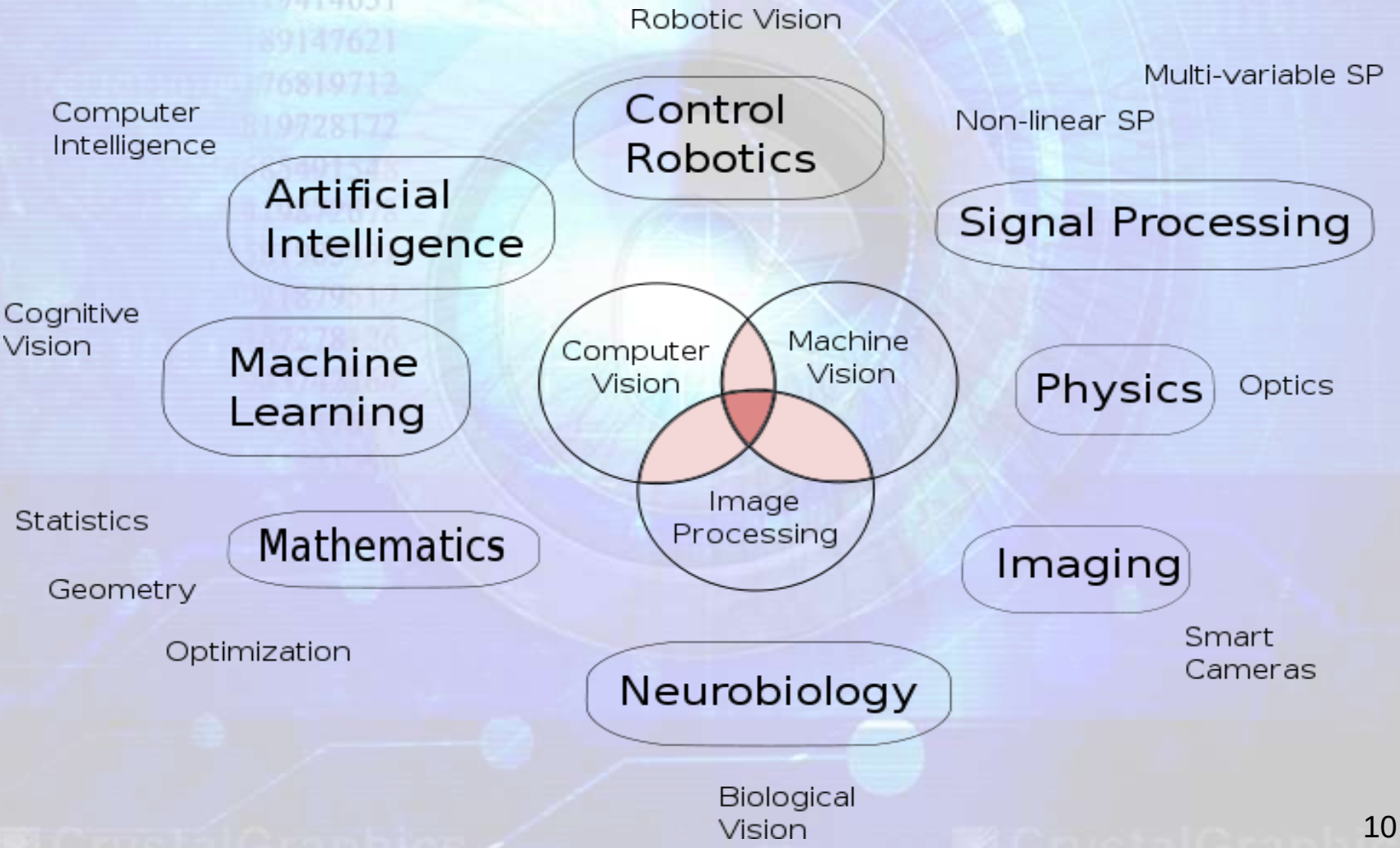

Comfort

Fun



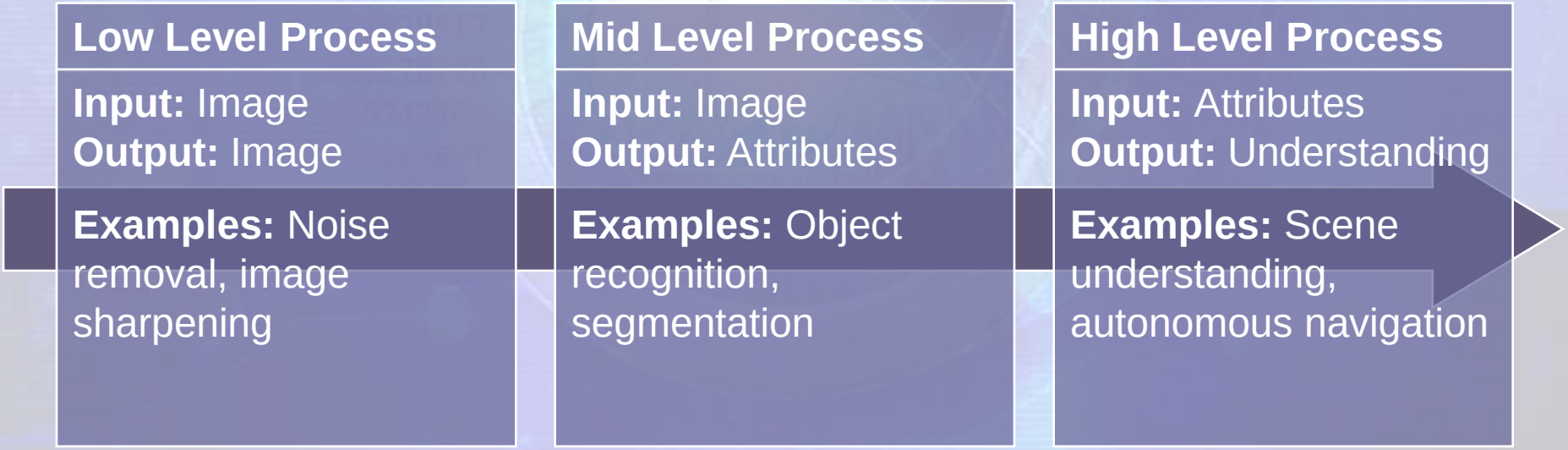

Access

# Connections to other Disciplines

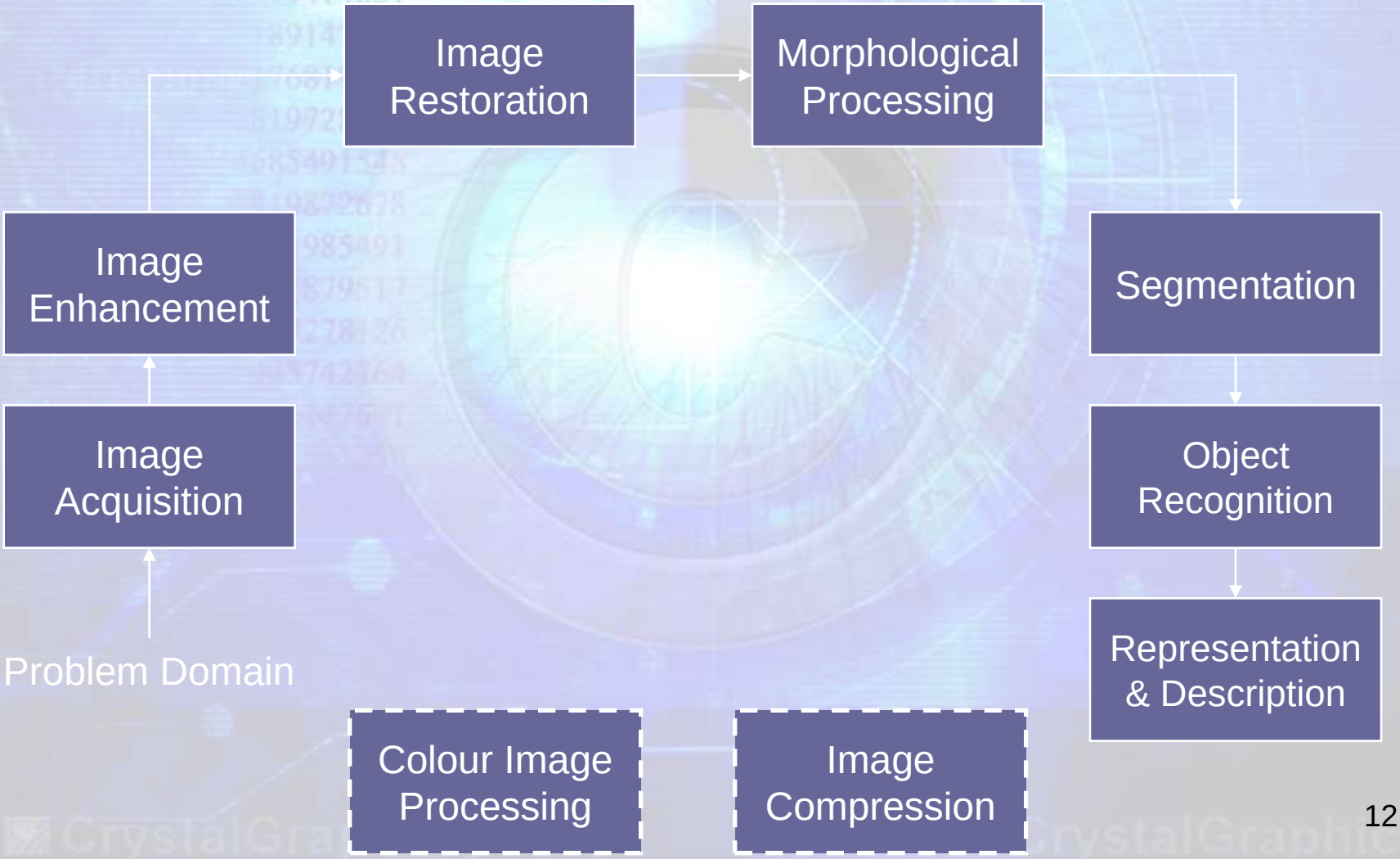


# What is Digital Image Processing?

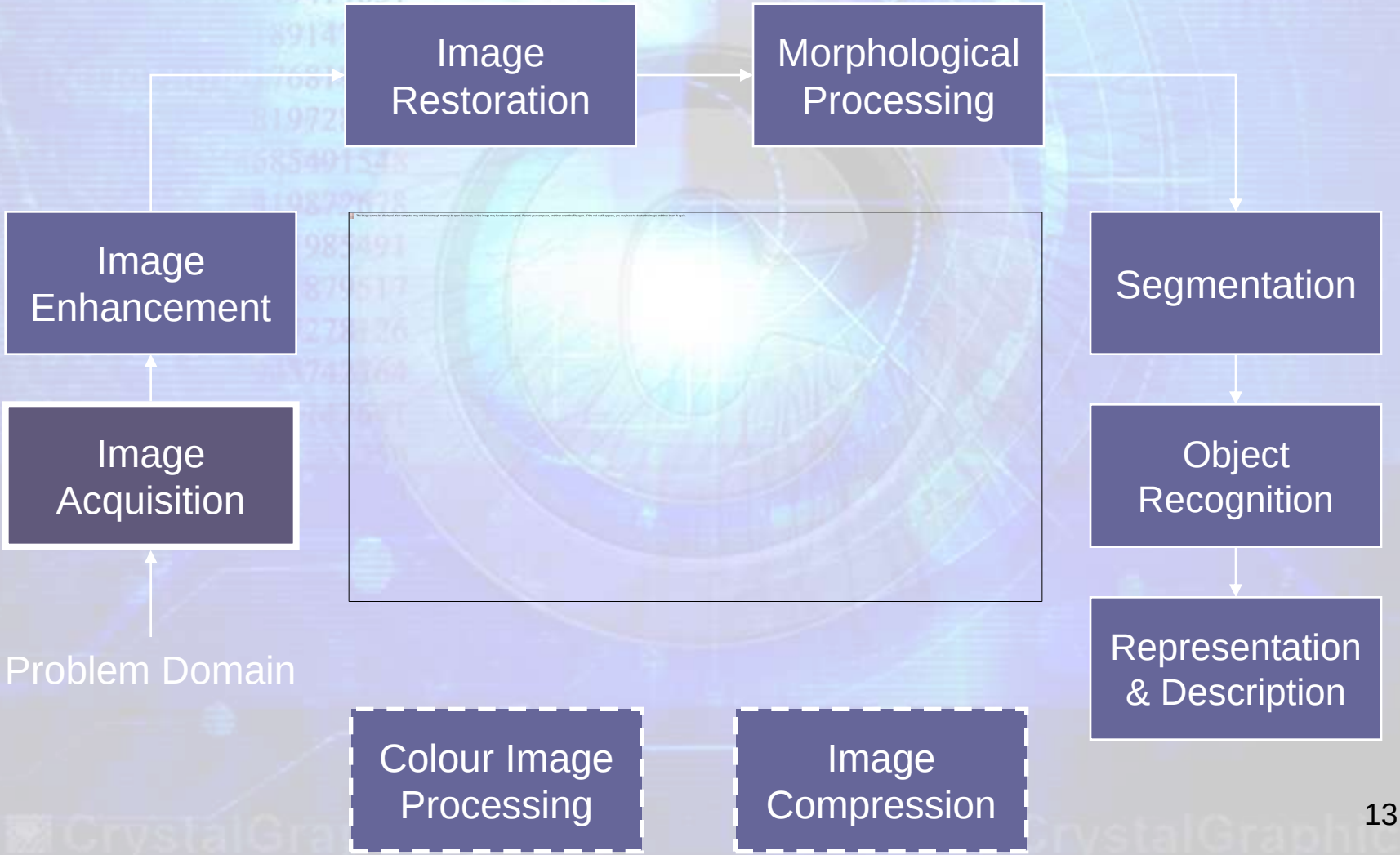
- The continuum from image processing to computer vision can be broken up into low-, mid- and high-level processes



# Key Stages in Digital Image Processing

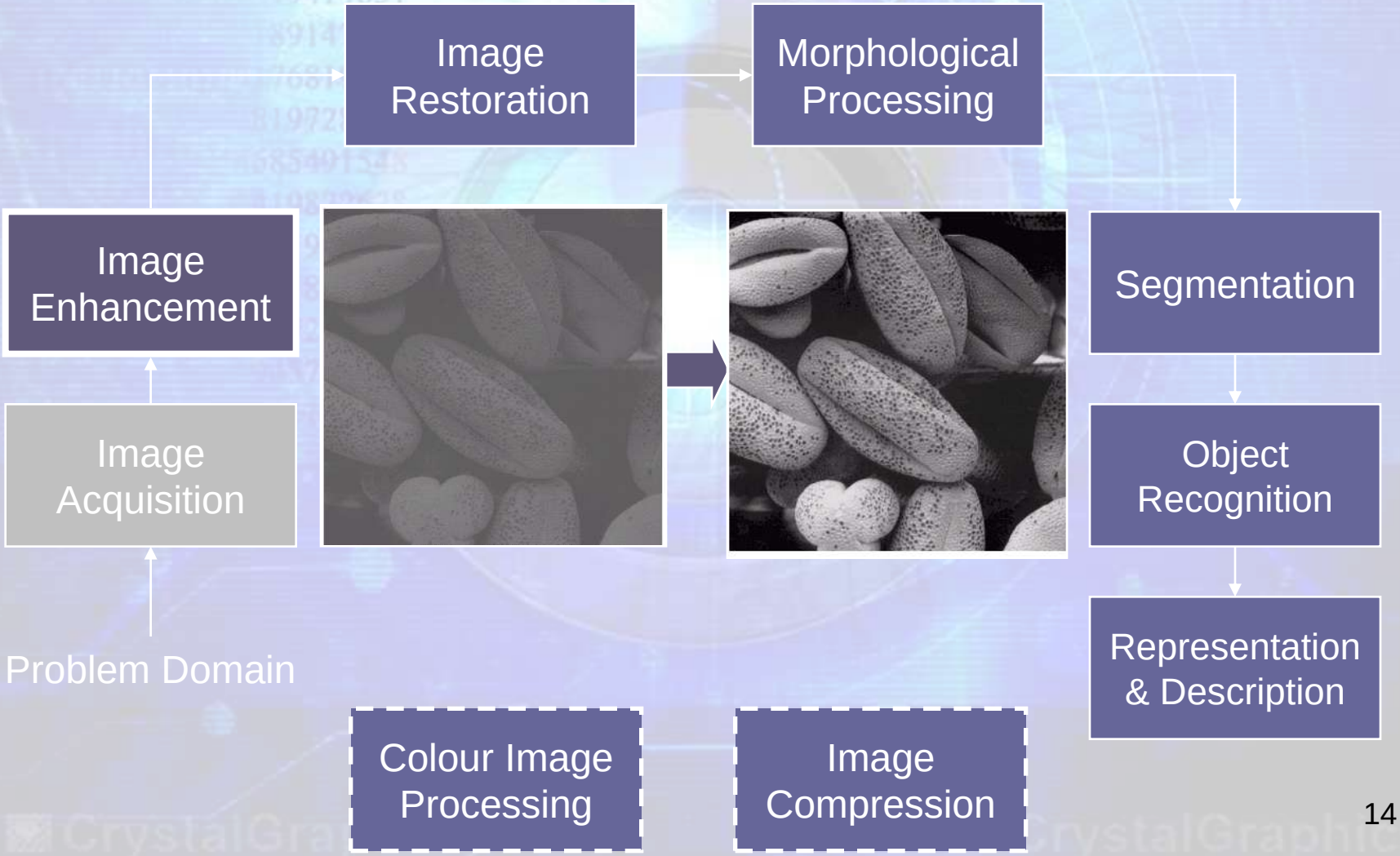


# Key Stages in Digital Image Processing: Image Acquisition



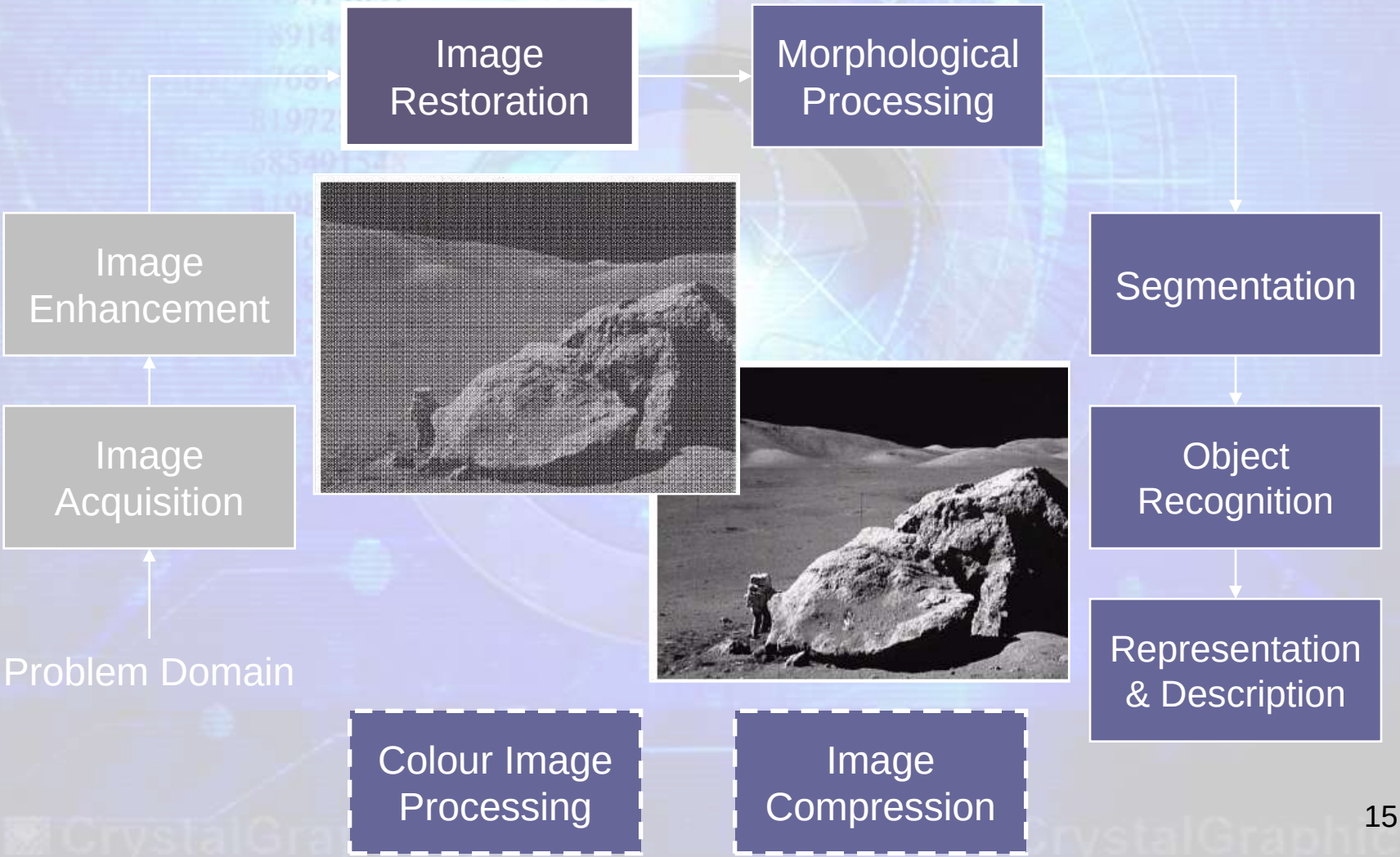

Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Image Enhancement



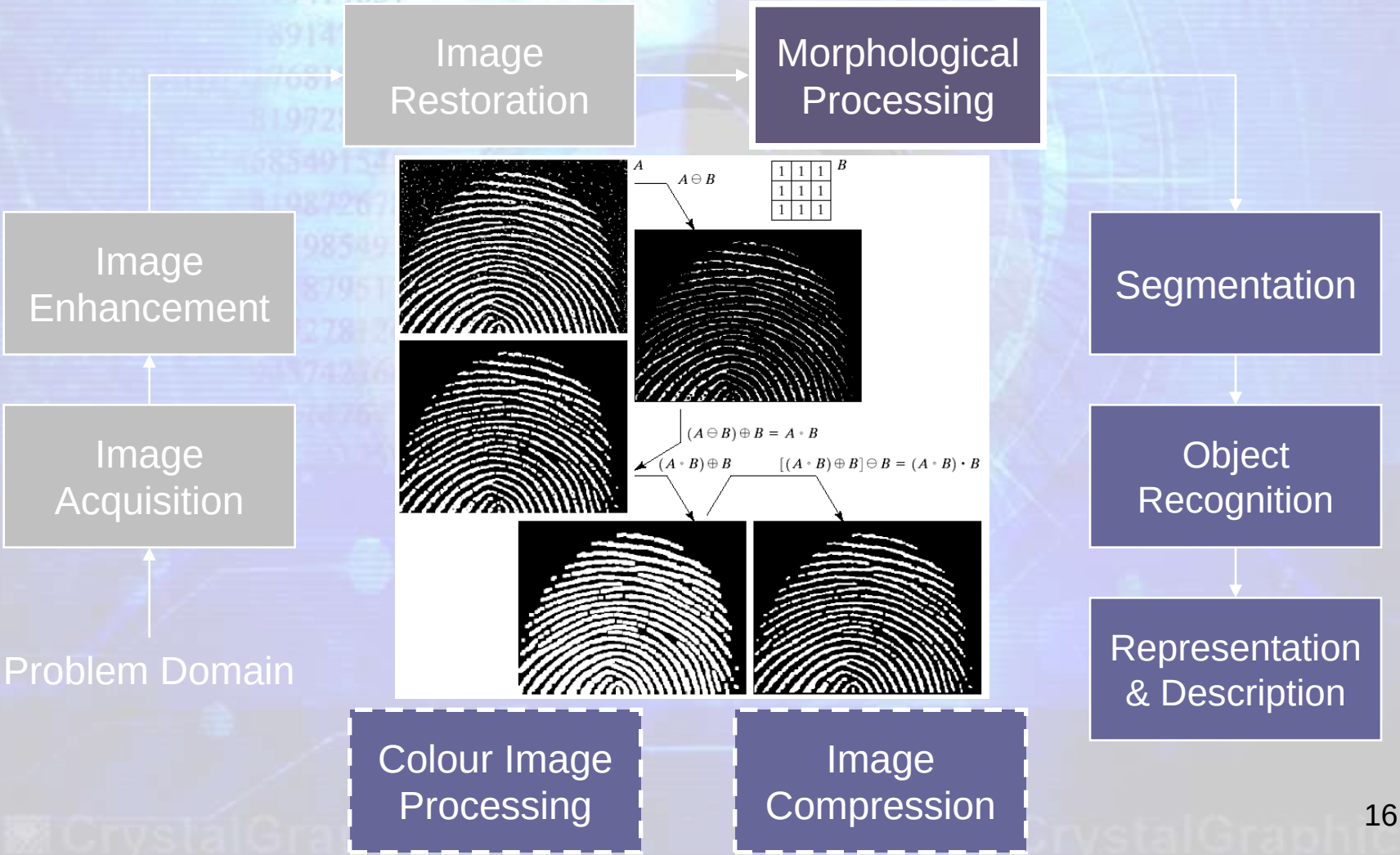

Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Image Restoration



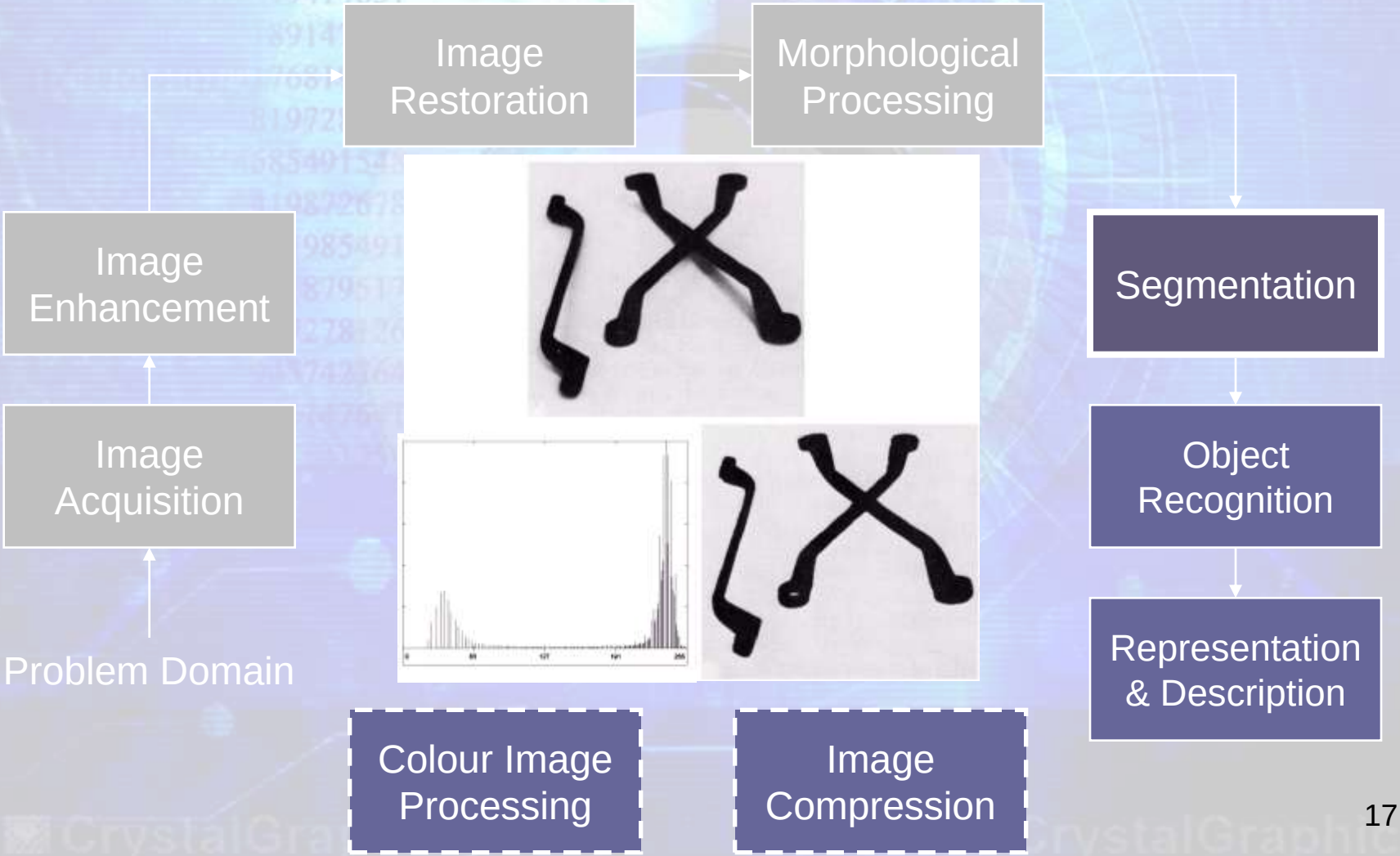

Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Morphological Processing




Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Segmentation




Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Object Recognition

Images taken from Gonzalez & Woods, Digital Image Processing (2002)

Problem Domain

Image Acquisition

Image Enhancement

Image Restoration

Morphological Processing

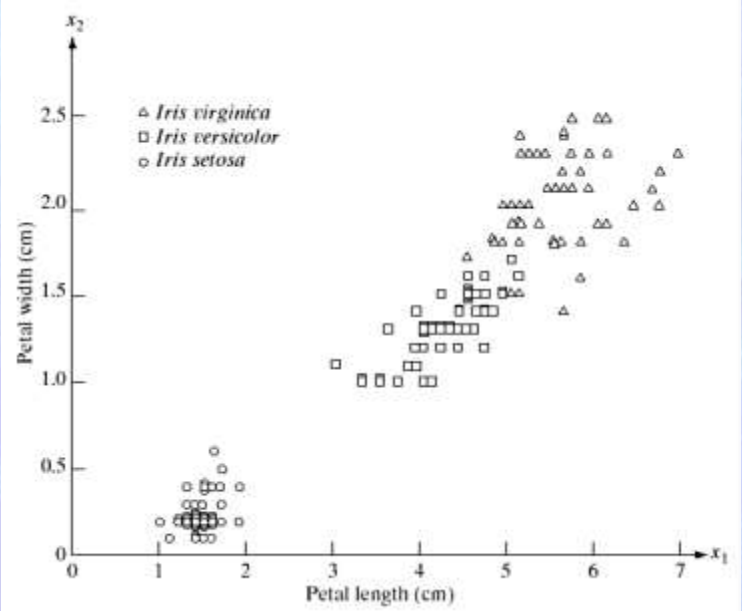
Segmentation

Object Recognition

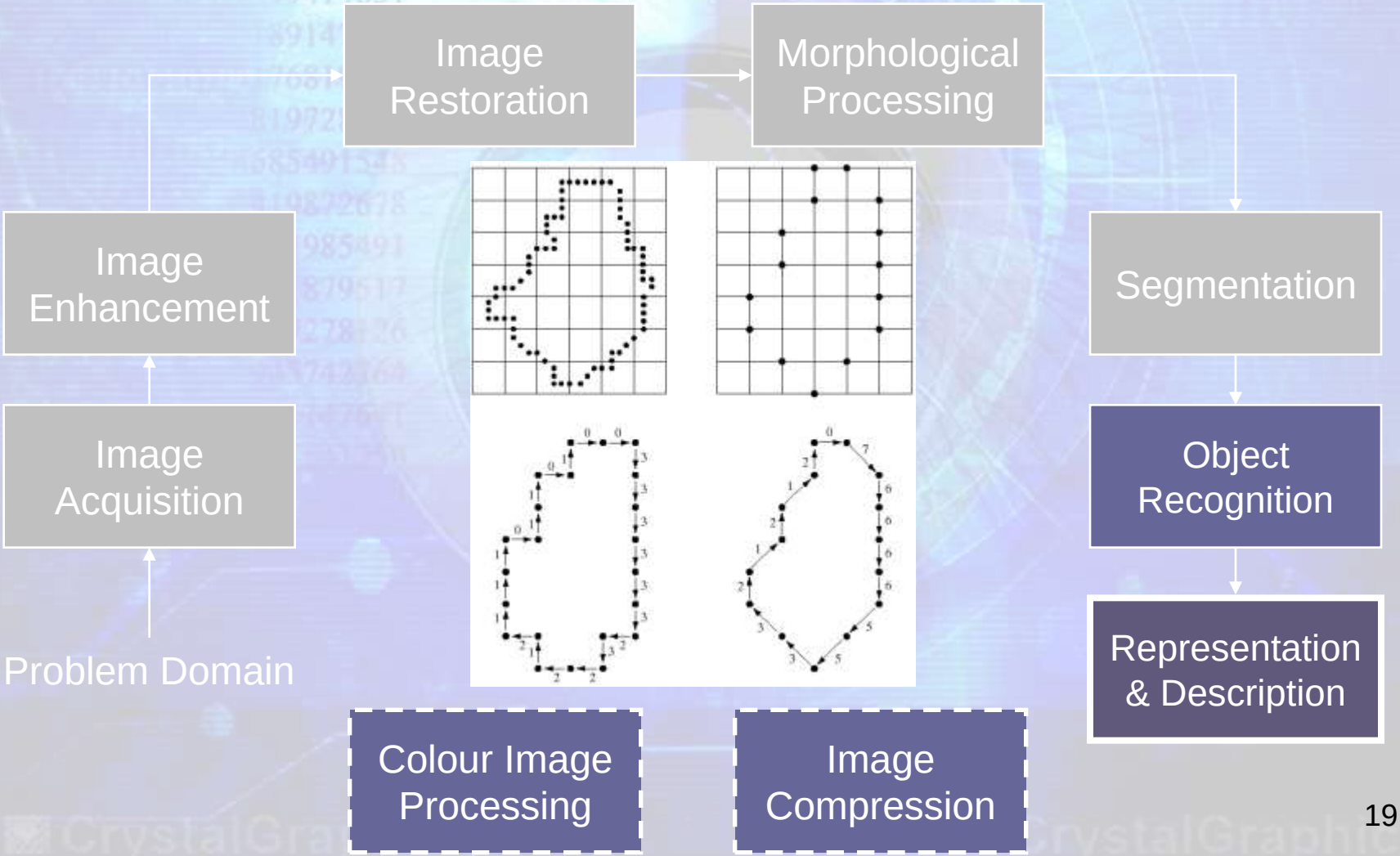
Representation & Description

Colour Image Processing

Image Compression



# Key Stages in Digital Image Processing: Representation & Description




Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# Key Stages in Digital Image Processing: Image Compression

Image Restoration

Morphological Processing

Segmentation

Object Recognition

Representation & Description

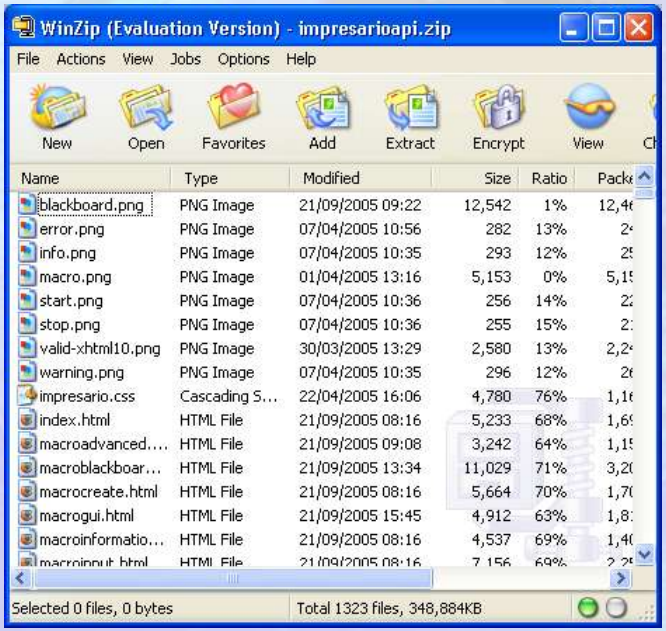
Image Compression

Colour Image Processing

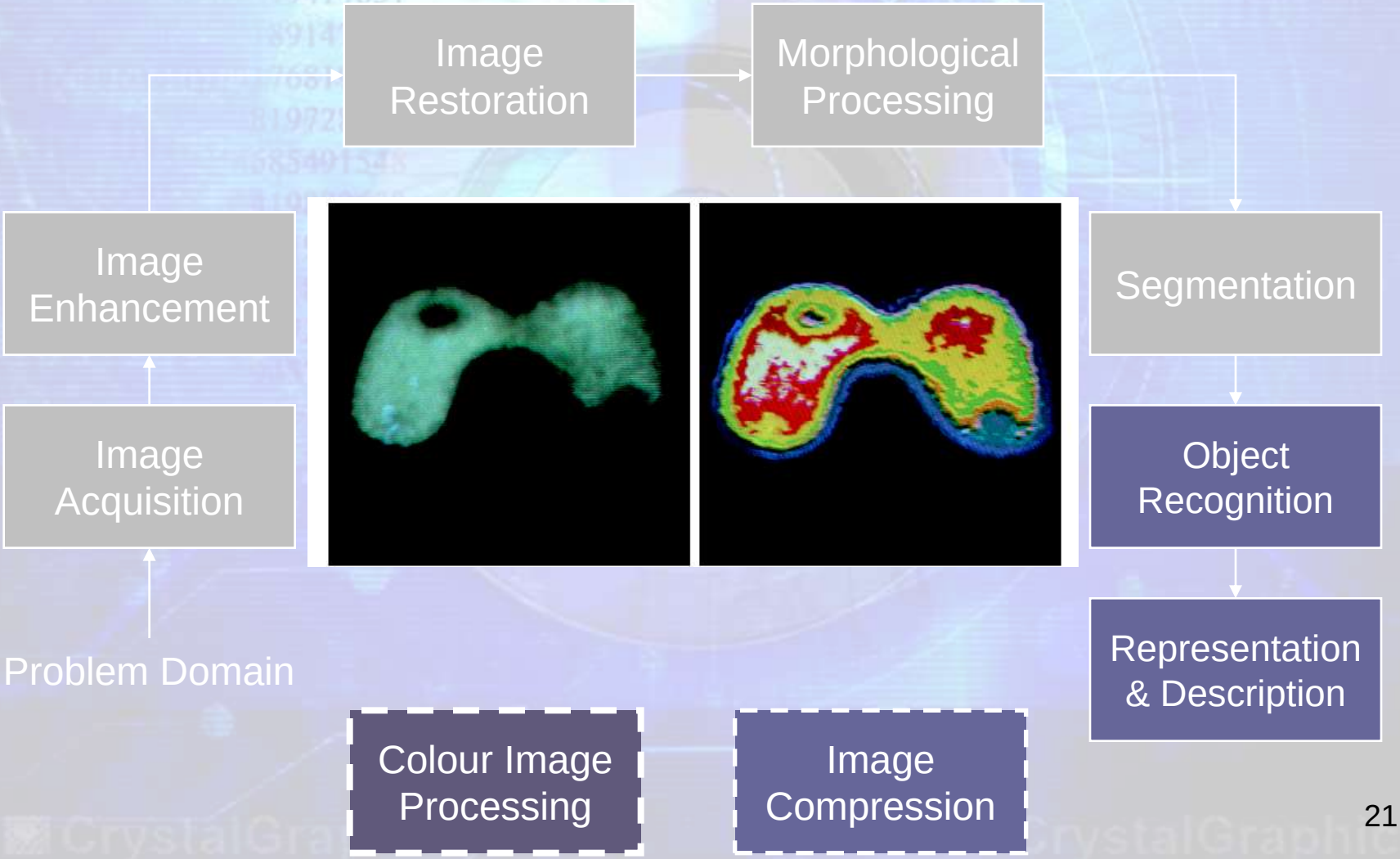
Image Enhancement

Image Acquisition

Problem Domain



# Key Stages in Digital Image Processing: Colour Image Processing




Images taken from Gonzalez & Woods, Digital Image Processing (2002)

# What is Object Recognition?

- Last step in image processing

- It is the task of finding and identifying objects in an image or video sequence

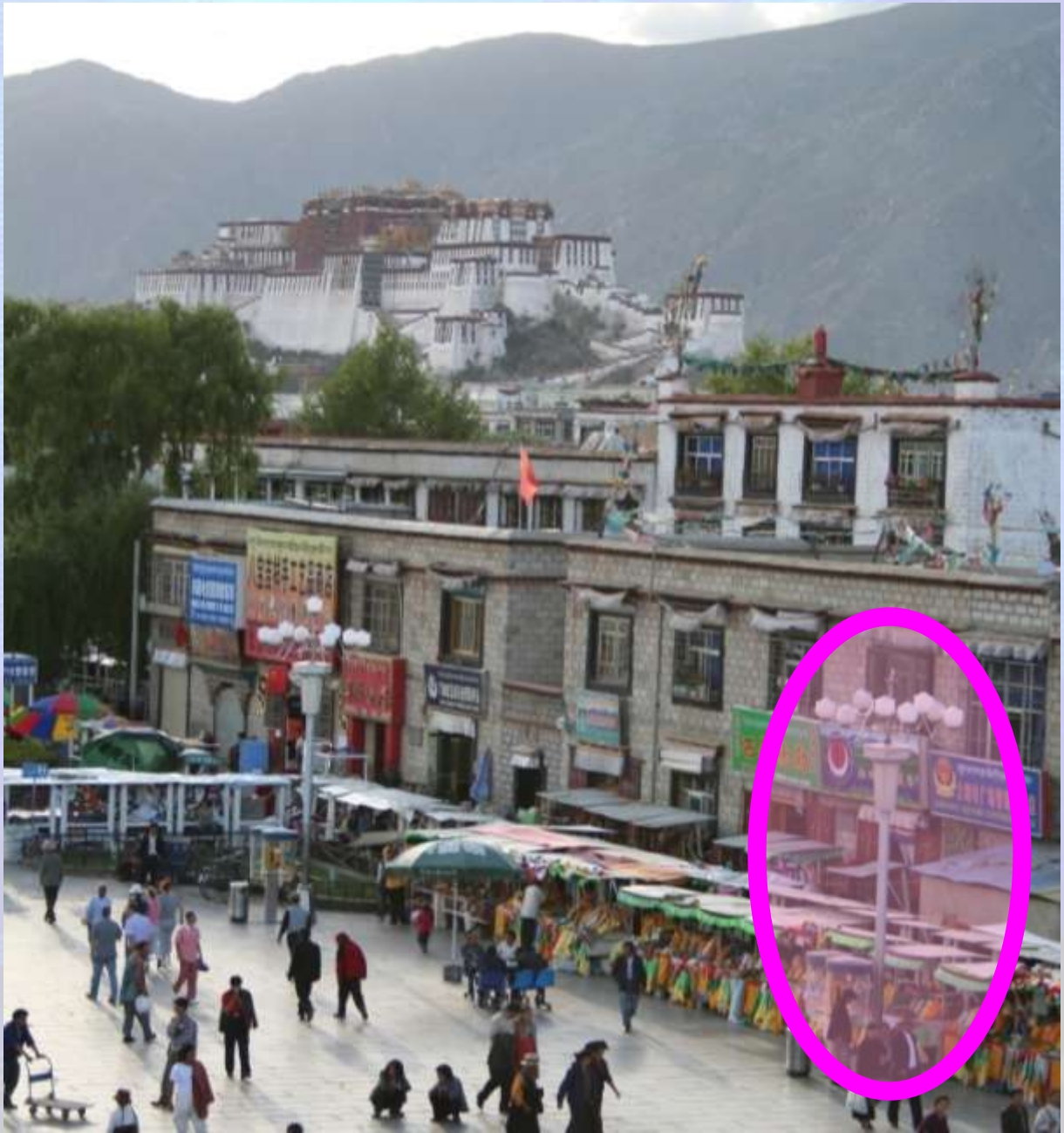
Like human understanding, it includes :

- ***Detection*** – of separate objects
- ***Description*** – of their geometry and positions in 3D
- ***Classification*** – as being one of a known class
- ***Identification*** – of the particular instance
- ***Understanding*** – of spatial relationships between objects

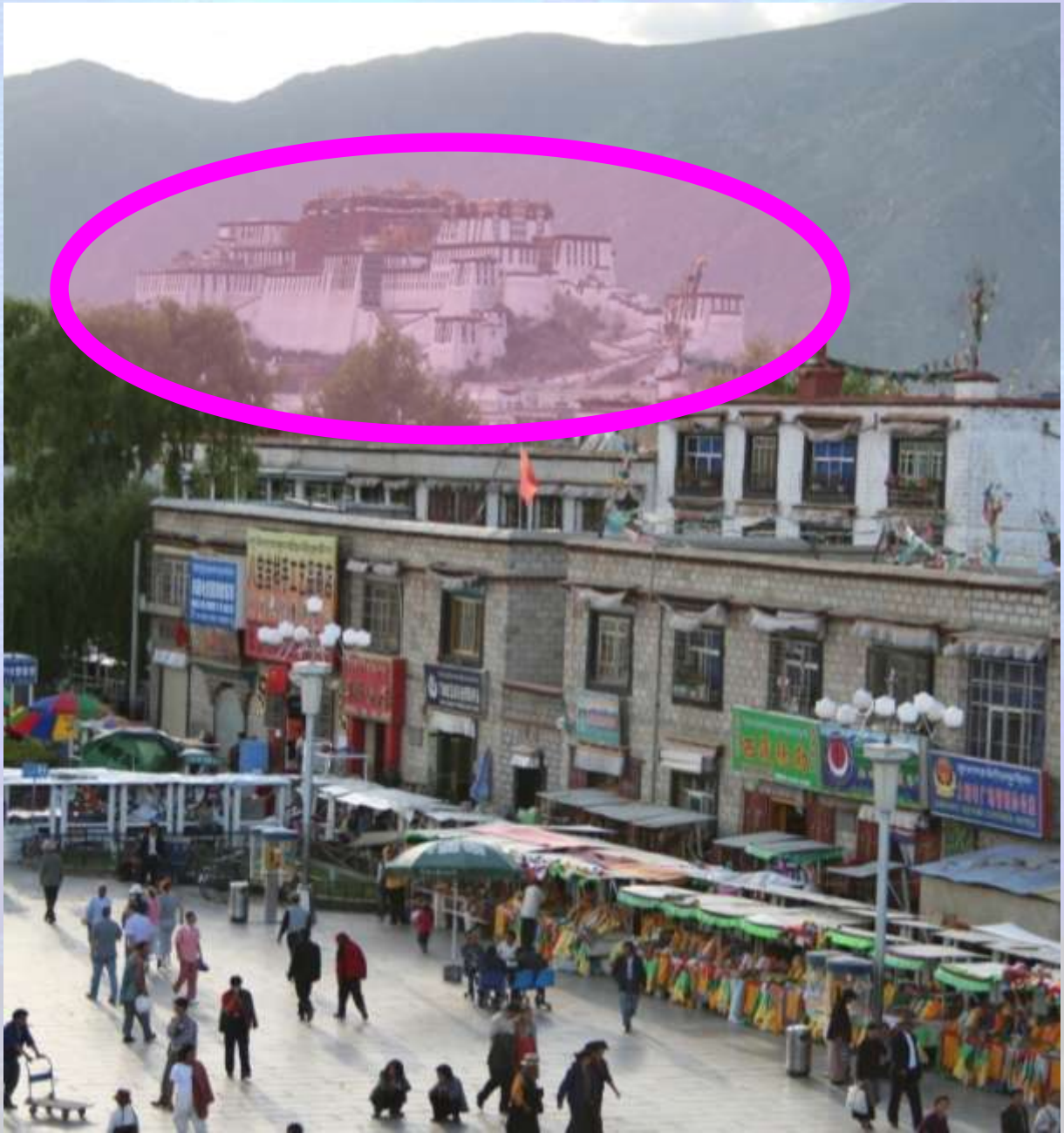
# So what does object recognition involve?

# Detection: are there people?

# Verification: is that a lamp?

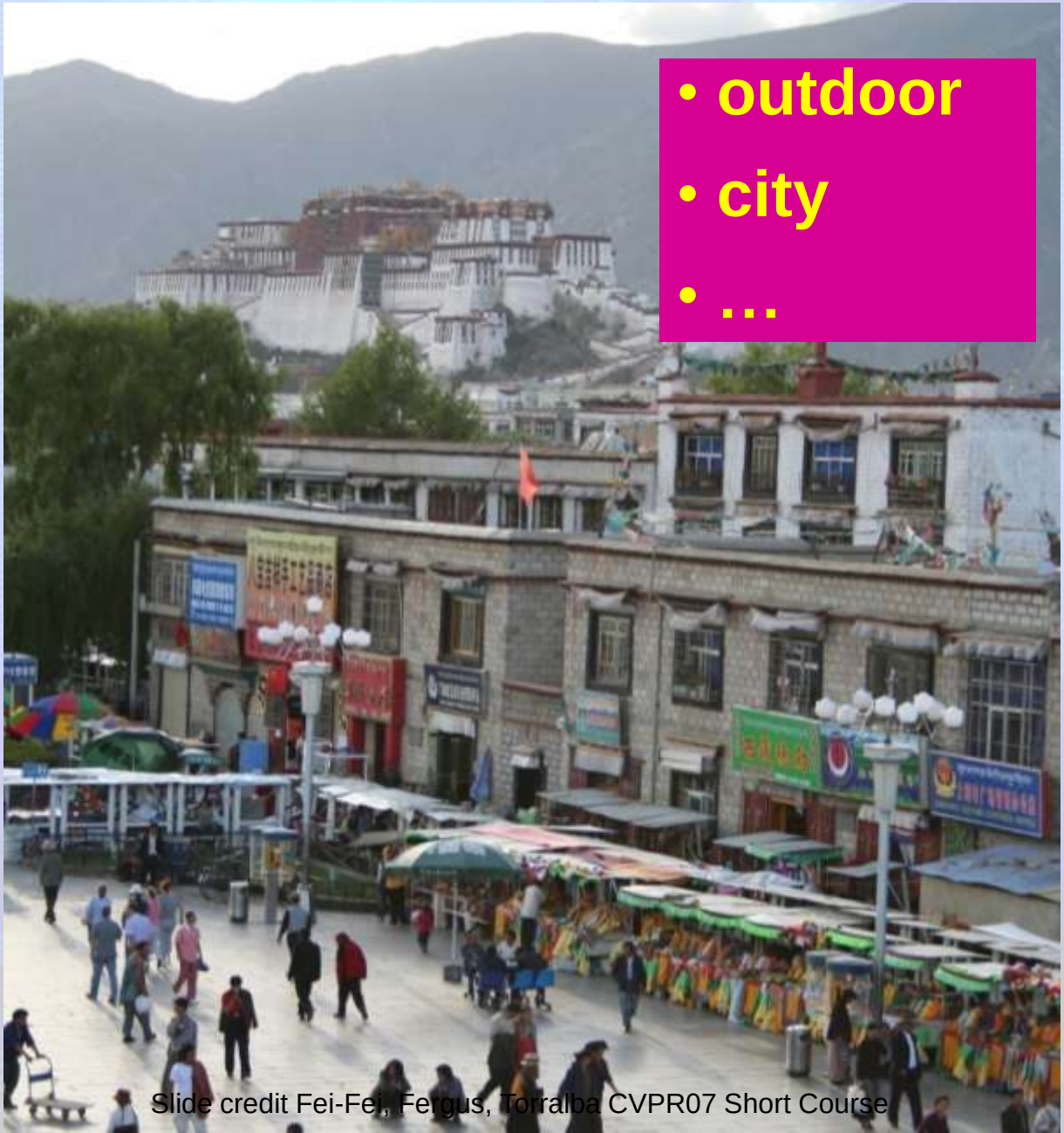


# Identification: is that Potala Palace?



# Object Categorization

# Scene and context categorization/Understanding



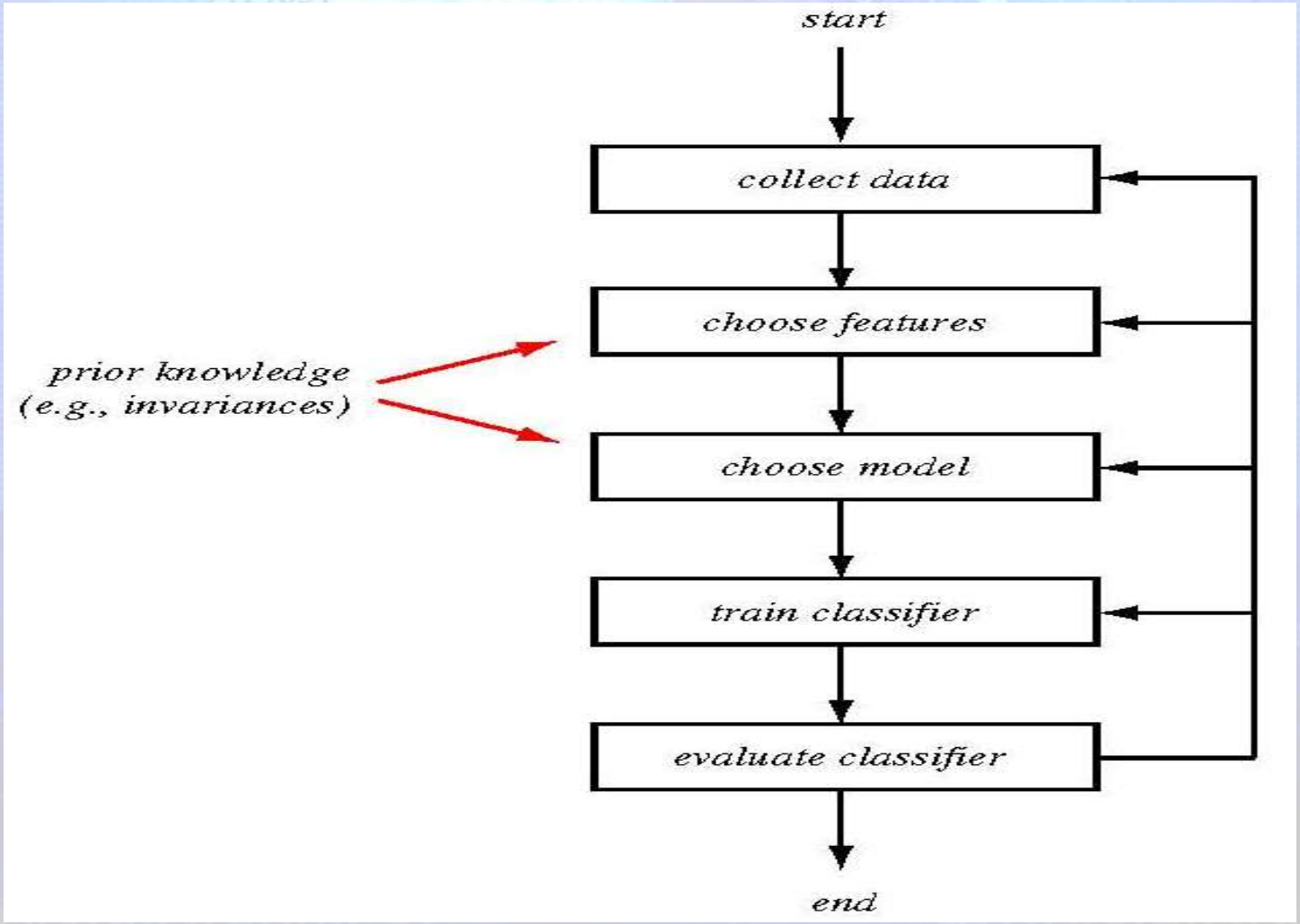

Slide credit Fei-Fei, Fergus, Torralba CVPR07 Short Course

# Learning and Adaptation

- ***Supervised learning***
  - A teacher provides a category label or cost for each pattern in the training set

- ***Unsupervised learning***
  - The system forms clusters or “natural groupings” of the input patterns

# The Design Cycle

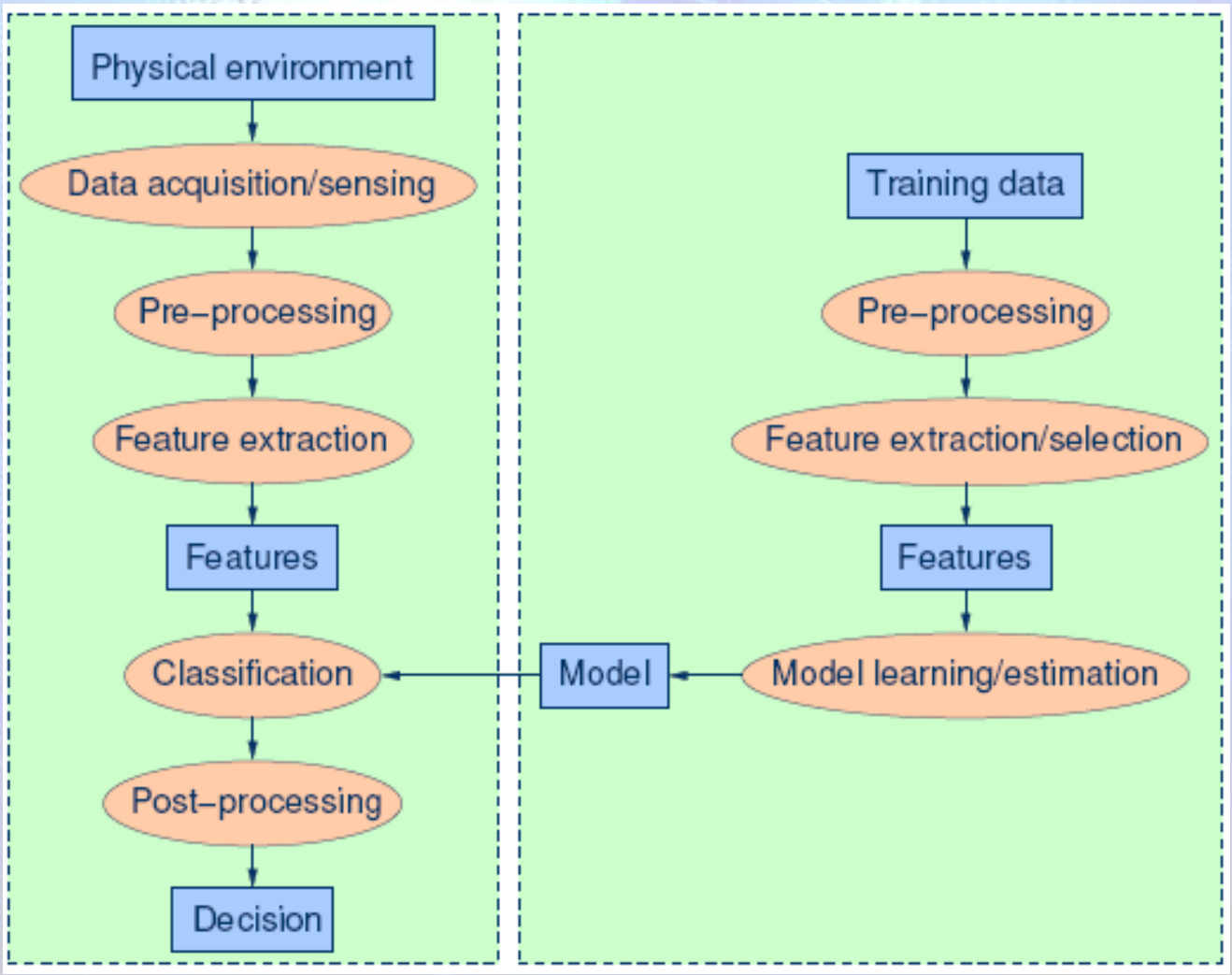


# What is Pattern Recognition?

**"The assignment of a physical object or event to one of several prespecified categeries" -- Duda & Hart**

- A ***pattern*** is an **object**, process or event that can be given a name.
- A ***pattern class*** (or category) is a set of patterns sharing common attributes and usually originating from the same source.
- During ***recognition*** (or ***classification***) given objects are assigned to prescribed classes.
- A ***classifier*** is a machine which performs classification.

# Basic Components of a Pattern Recognition System



# Components of Pattern Recognition (Cont'd)

- **Data acquisition and sensing**
- **Pre-processing**
  - Removal of noise in data.
  - Isolation of patterns of interest from the background.
- **Feature extraction**
  - Finding a new representation in terms of features.

  (Better for further processing)

# Components of Pattern Recognition (Cont'd)

- **Model learning and estimation**
  - Learning a mapping between features and pattern groups.
- **Classification**
  - Using learned models to assign a pattern to a predefined category
- **Post-processing**
  - Evaluation of confidence in decisions.
  - Exploitation of context to improve performances.

# Examples of Pattern Recognition Applications

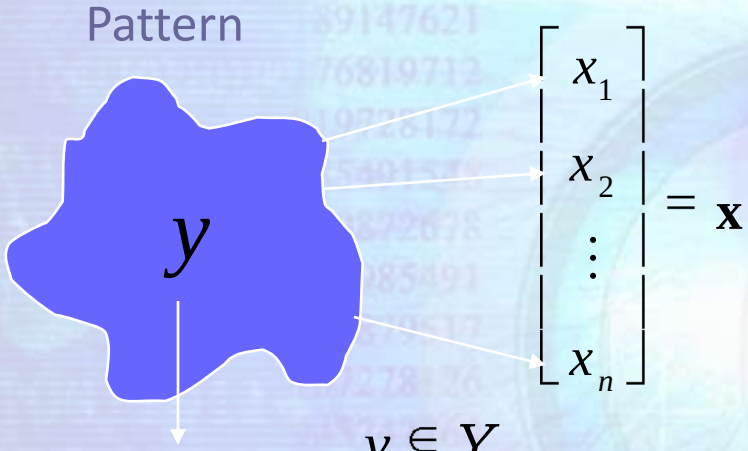
| Problem Domain | Application | Input Pattern | Pattern Classes |
|---|---|---|---|
| Document image analysis | Optical character recognition | Document image | Characters, words |
| Document classification | Internet search | Text document | Semantic categories |
| Document classification | Junk mail filtering | Email | Junk/non-junk |
| Multimedia database retrieval | Internet search | Video clip | Video genres |
| Speech recognition | Telephone directory assistance | Speech waveform | Spoken words |
| Natural language processing | Information extraction | Sentences | Parts of speech |
| Biometric recognition | Personal identification | Face, iris, fingerprint | Authorized users for access control |
| Medical | Computer aided diagnosis | Microscopic image | Cancerous/healthy cell |
| Military | Automatic target recognition | Optical or infrared image | Target type |
| Industrial automation | Printed circuit board inspection | Intensity or range image | Defective/non-defective product |
| Industrial automation | Fruit sorting | Images taken on a conveyor belt | Grade of quality |
| Remote sensing | Forecasting crop yield | Multispectral image | Land use categories |
| Bioinformatics | Sequence analysis | DNA sequence | Known types of genes |
| Data mining | Searching for meaningful patterns | Points in multidimensional space | Compact and well-separated clusters |

# Pattern Representation

- A pattern is represented by a set of *d* features, or attributes, viewed as a *d*-dimensional *feature vector*.

$$\mathbf{x} = (x_1, x_2, \cdots, x_d)^T$$

# Basic concepts




**Feature vector** $\mathbf{x} \in X$

- A vector of observations (measurements).
- $\mathbf{x}$ is a point in feature space $X$

**Hidden state**

- Cannot be directly measured.
- Patterns with equal hidden state belong to the same class.

**Task**

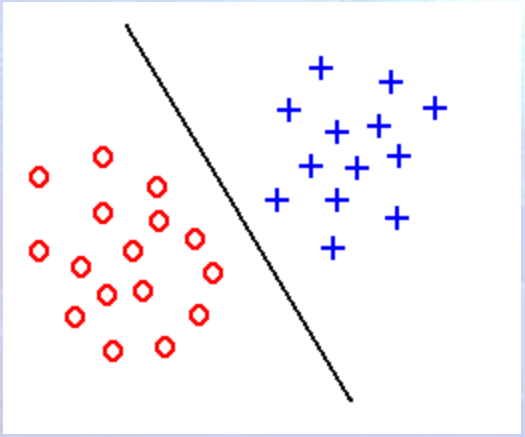
- To design a classifer (decision rule) $q : X \to Y$

which decides about a hidden state based on an onbservation.

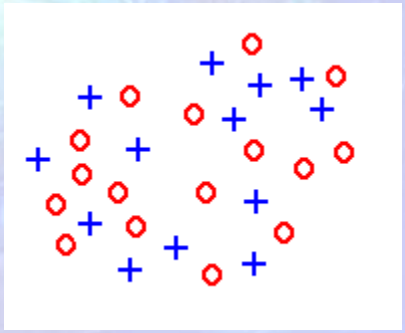
# Feature Extraction

**Task:** to extract features which are good for classification.

Good features:
- Objects from the same class have similar feature values.
- Objects from different classes have different values.



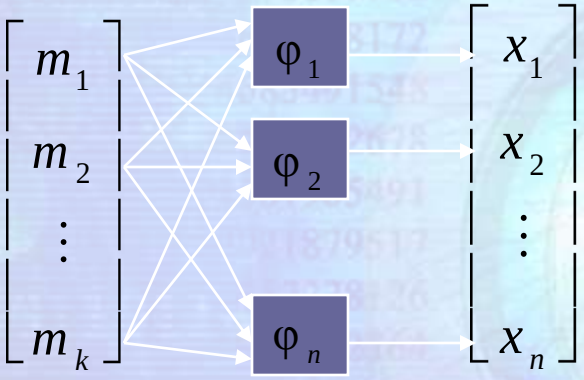
"Good" features



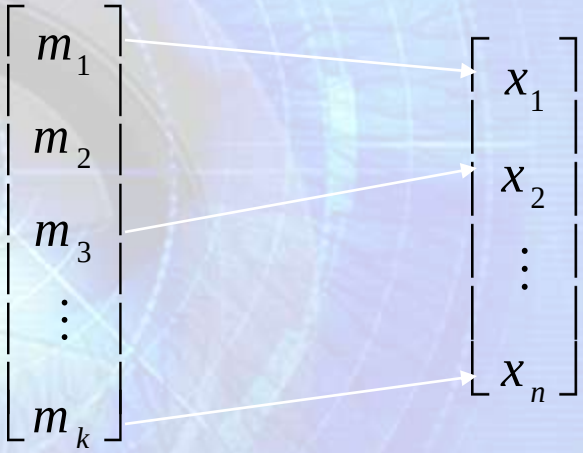
"Bad" features

# Feature Extraction Methods

Feature extraction




Feature selection




Problem can be expressed as optimization of parameters of featrure extractor $\varphi(\theta)$

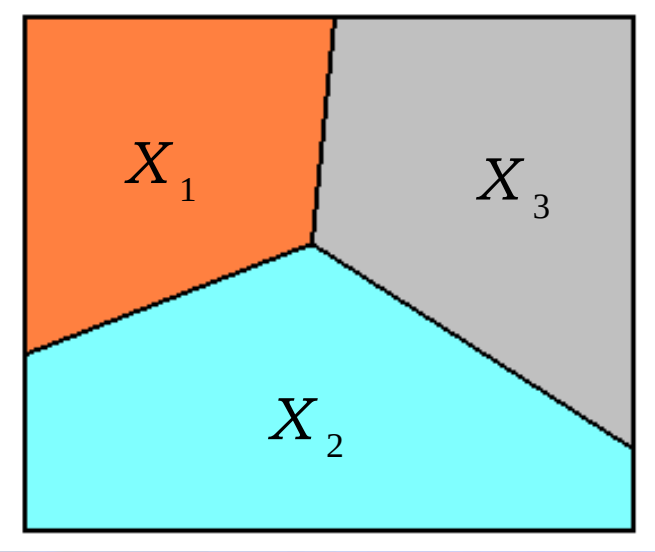
**Supervised methods**: objective function is a criterion of separability (discriminability) of labeled examples, e.g., linear discriminat analysis (LDA).

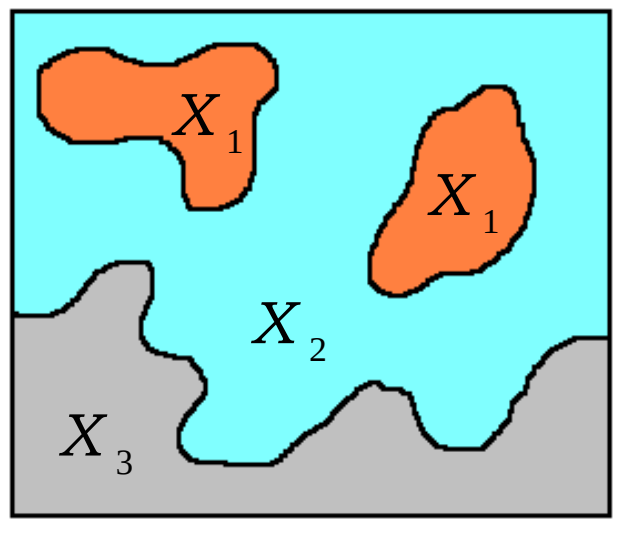
**Unsupervised methods**: lower dimesional representation which preserves important characteristics of input data is sought for, e.g., principal component analysis (PCA).

# Classifier

A classifier partitions feature space *X* into **class-labeled regions** such that

$$X = X_1 \cup X_2 \cup \ldots \cup X_{|Y|} \quad \text{and} \quad X_1 \cap X_2 \cap \ldots \cap X_{|Y|} = \{0\}$$




The classification consists of determining to which region a feature vector **x** belongs to.

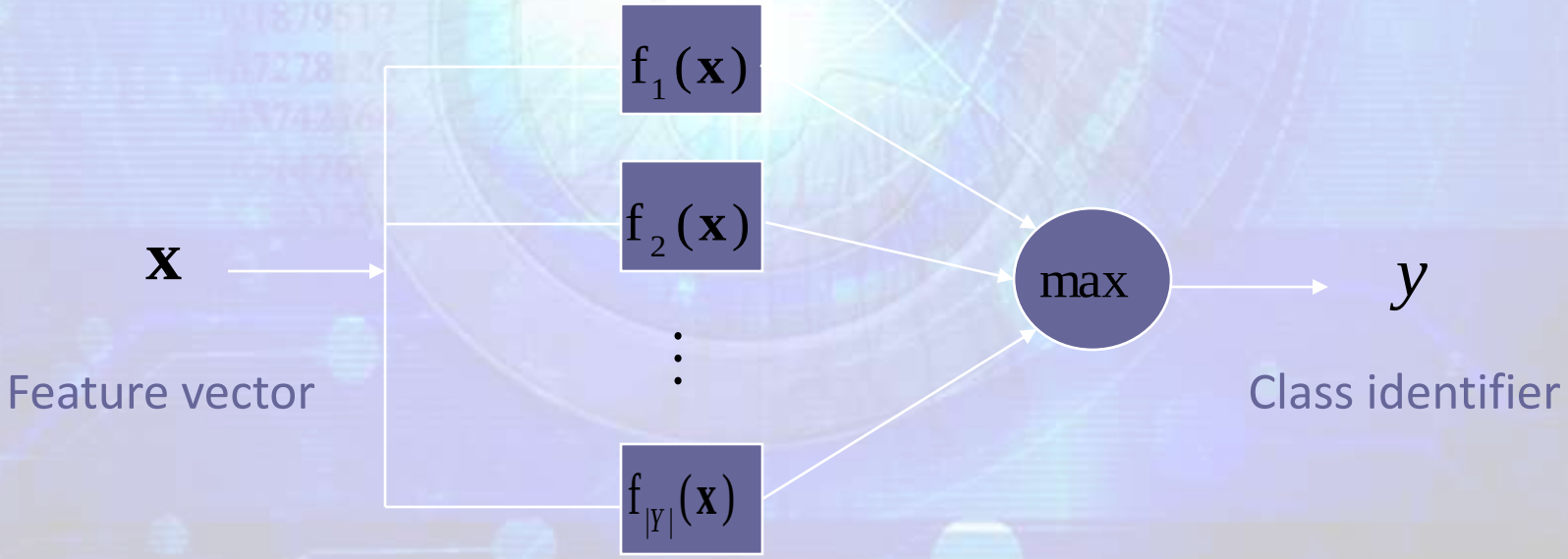
Borders between **decision boundaries** are called decision regions.

# Representation of classifier

A classifier is typically represented as a set of discriminant functions

$$\mathrm{f}_i(\mathbf{x}) : X \rightarrow \Re, i = 1, \ldots, |Y|$$

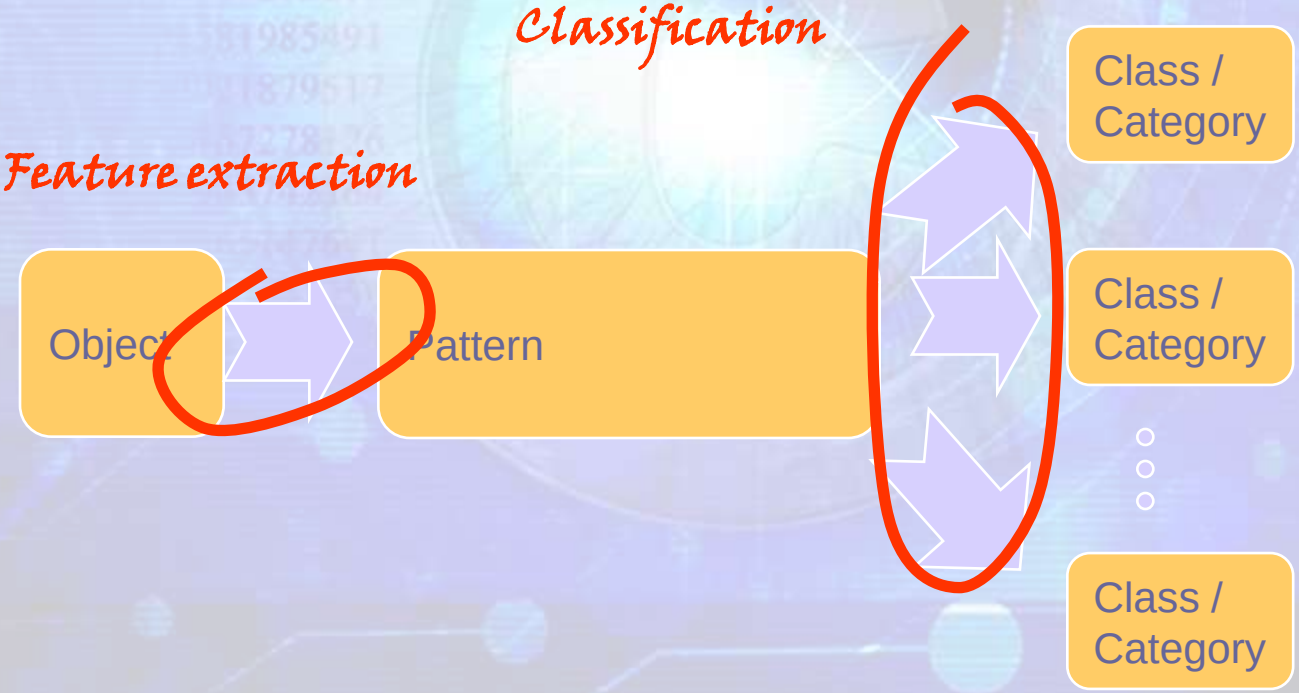
The classifier assigns a feature vector **x** to the *i*-the class if $\mathrm{f}_i(\mathbf{x}) > \mathrm{f}_j(\mathbf{x}) \quad \forall j \neq i$



# Block diagram

Both definitions may be depicted by the following block diagram.

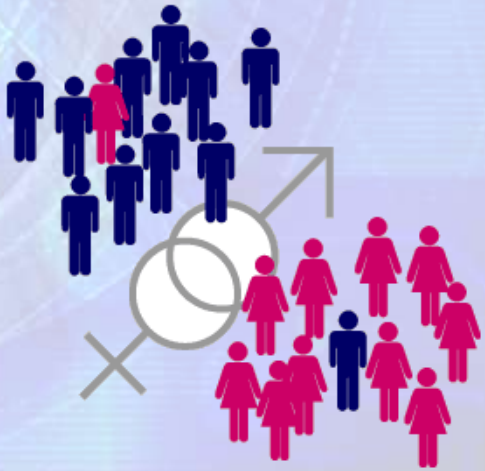
The process consists of two major operations:



# Example : Gender

Assume an algorithm to recognize the gender of a student in a university, where the available input is several features of the students (of course, the gender cannot be one of the features).

The student to be classified is the ***Object***, The gender (Male or Female) are the ***Classes***, and the input which is referred to the student is the *Pattern*.



# What is a *Feature*?

***Feature*** is a scalar *x* which is quantitatively describes a property of the ***Object***.

Example: Possible features of a student:

- Number of eyes — $x \in \{0, 1, 2\}$
- Hair color — $x \in \{0$ =Black, $1$ =Blond, $2$ =Red $,\ldots\}$
- Wear glasses or not — $x \in \{0, 1\}$
- Hair length [cm] — $x \in [0..100]$
- Shoe size [u.s] — $x \in [3.5, 4, 4.5, .. ,14]$
- Height [cm] — $x \in [40..240]$
- Weight [kg] — $x \in [30..600]$

# What is *Feature Extraction*?

"When we have two or more classes, feature extraction consist of choosing those features which are most effective for preserving class separability" (Fukunaga p. 441)

Assume we choose the shoe size of the student as a feature. The selection is heuristically and seems reasonable.

In this case, asking the student to raise his leg and reading the size of the show is ***Feature extraction***…

# What is a *Pattern*?

***Pattern*** is a *n*-tuple $X$ (vector) of $N$ scalars $x_i$ $i \in [1,N]$, which are called the ***Features***.

Conventional form of a pattern is:

$$X = \begin{matrix} x_1 & x_2 & \cdots & x_N \end{matrix}^T, \qquad x_i \in \square, \; X \in V$$

Where $V$ is known as the ***Feature Space***, and $N$ is the dimension of $V$.

# Possible patterns for the gender problem:

**We can use the shoe size alone to have:**

$$X = Shoe\ size$$

**We can combine the height and the weight to have:**

$$X = height, weight^{T}$$

**We can even combine the height, weight and the shoe size to be on the safe side:**

$$X = height, weight, shoe\ size^{T}$$

**Or, we can use them all:**

$$X = \#\ of\ eyes, hair\ color, glasses, hair\ length, height, weight, shoe\ size^{T}$$

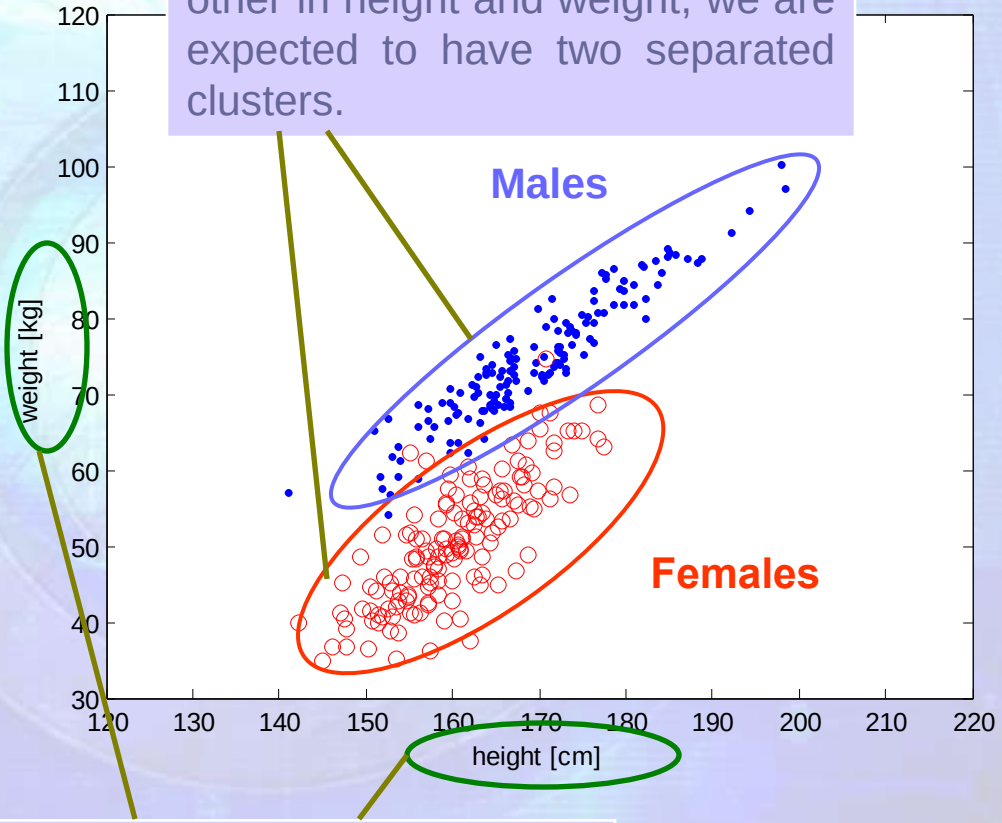
# Example

Assume we are using the height and the weight of each of the students in the university as a ***pattern***.

The height and the weight are both *features*, which span a *feature space V* of dimension 2.

Each of the students is represented as a point in the ***feature space***. Patterns of male students are depicted in blue, and those of female students – in red.

Since the male students and the female students differ from each other in height and weight, we are expected to have two separated clusters.




Each student is characterized by a vector of two feature: (height, weight).

# What is a *Class*?

"Class is a set of ***patterns*** that share some common properties" (Wang p.10)

In our example, the **Male students** and the **Female students** are two classes of objects that share a common gender.

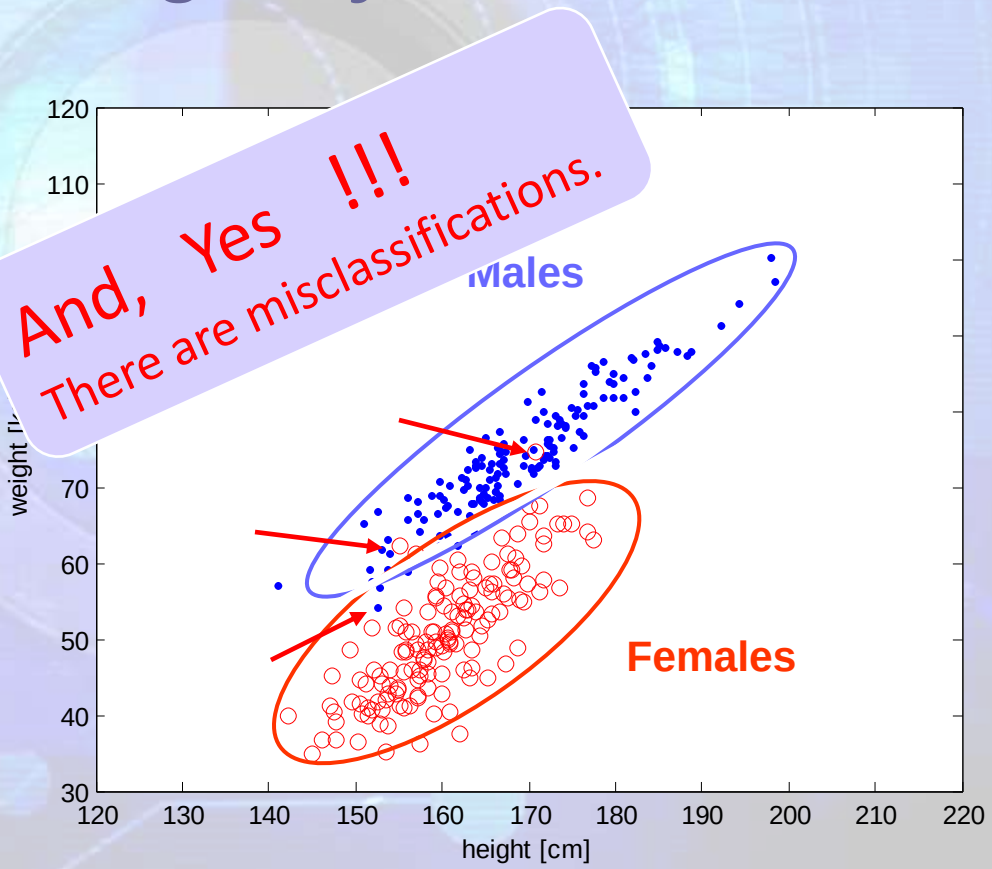
*BUT, pay attention to the fact that two different classes do not imply two separated clusters in the feature space. In fact, in the gender example, we can find at least three female points in the male cluster.*

# What is Classification?

***Classification*** is a mathematical function or algorithm which assigns a ***feature*** to one of the ***classes***.

Example:

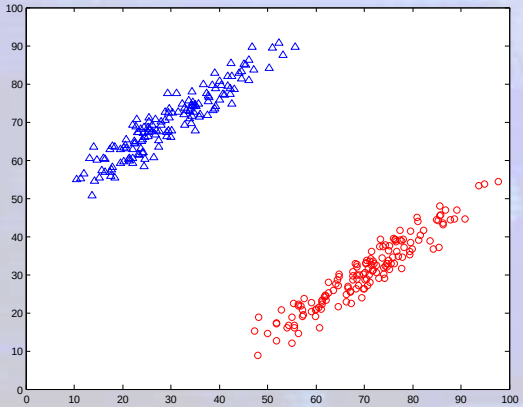
We can draw a line between the two clusters in the gender example, and every student will be classified as a female or male according to this line.



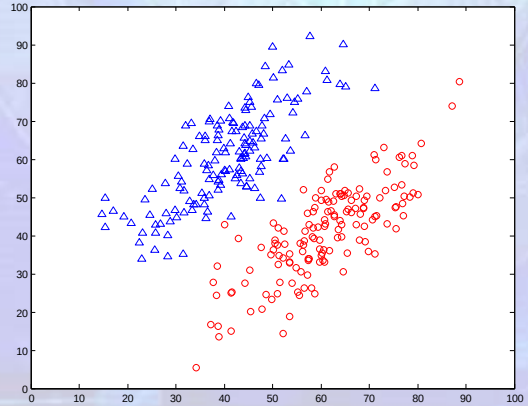
# Clusters Separation

Misclassifications are a consequence of the separation of the clusters. The separation of clusters is quantified using two major methods:

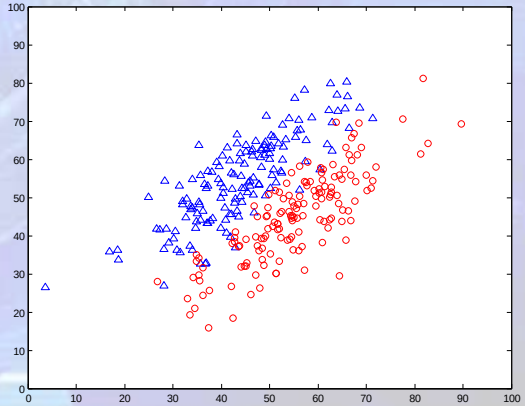
1. Mathematically: there are several *separation criteria*'s.
2. "Intuitively": overlapping of the clusters.



Separable clusters



Almost separable clusters



Non-separable clusters

# Classification Quality

WARNING!!! Although the idea is well illustrated, it is a bad habit to judge a classification quality according to the visual representation of clusters.

The classification quality is strongly depends on the clusters separation

The clusters separation strongly depends on the ***features selection***

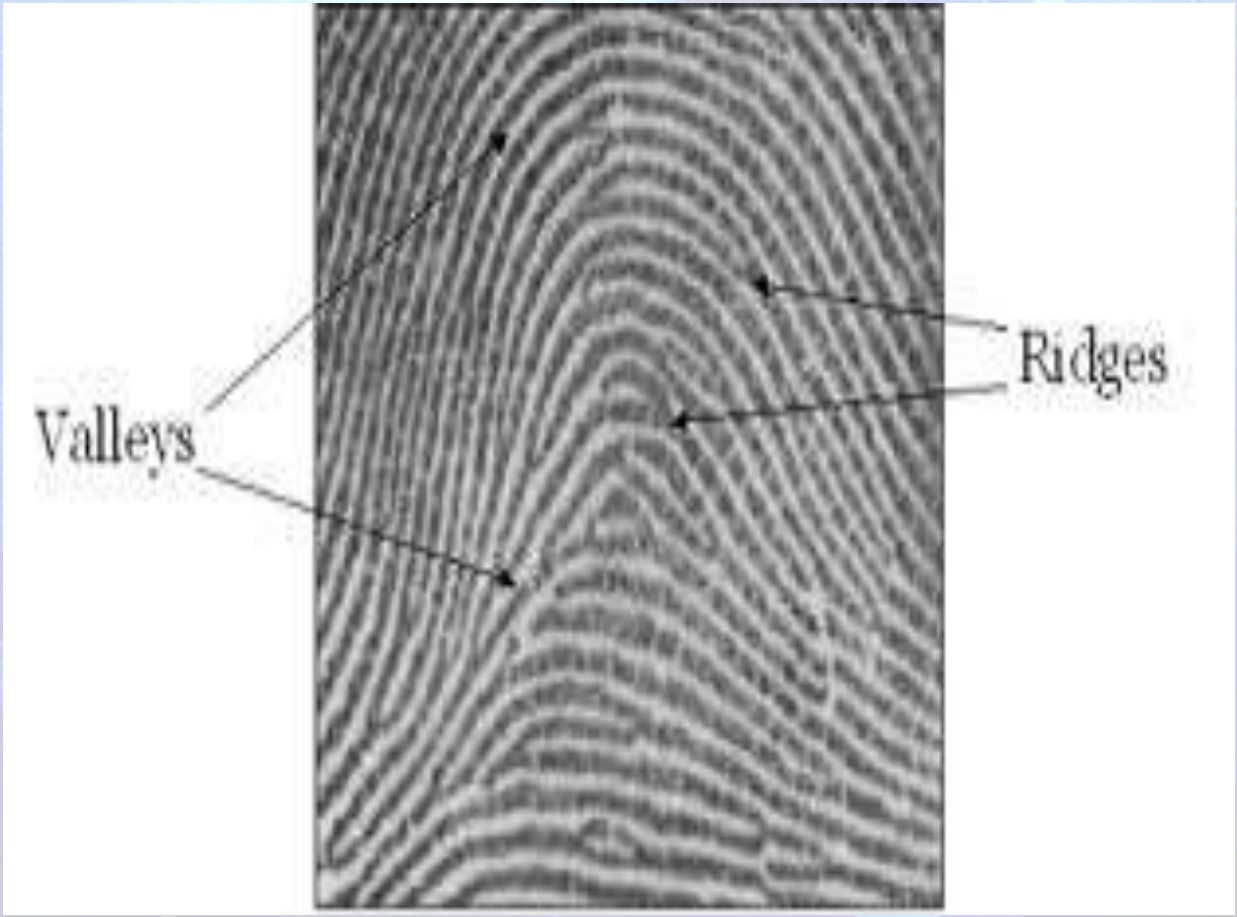
***Feature selection*** is of paramount importance in classification quality

# Fingerprint Recognition

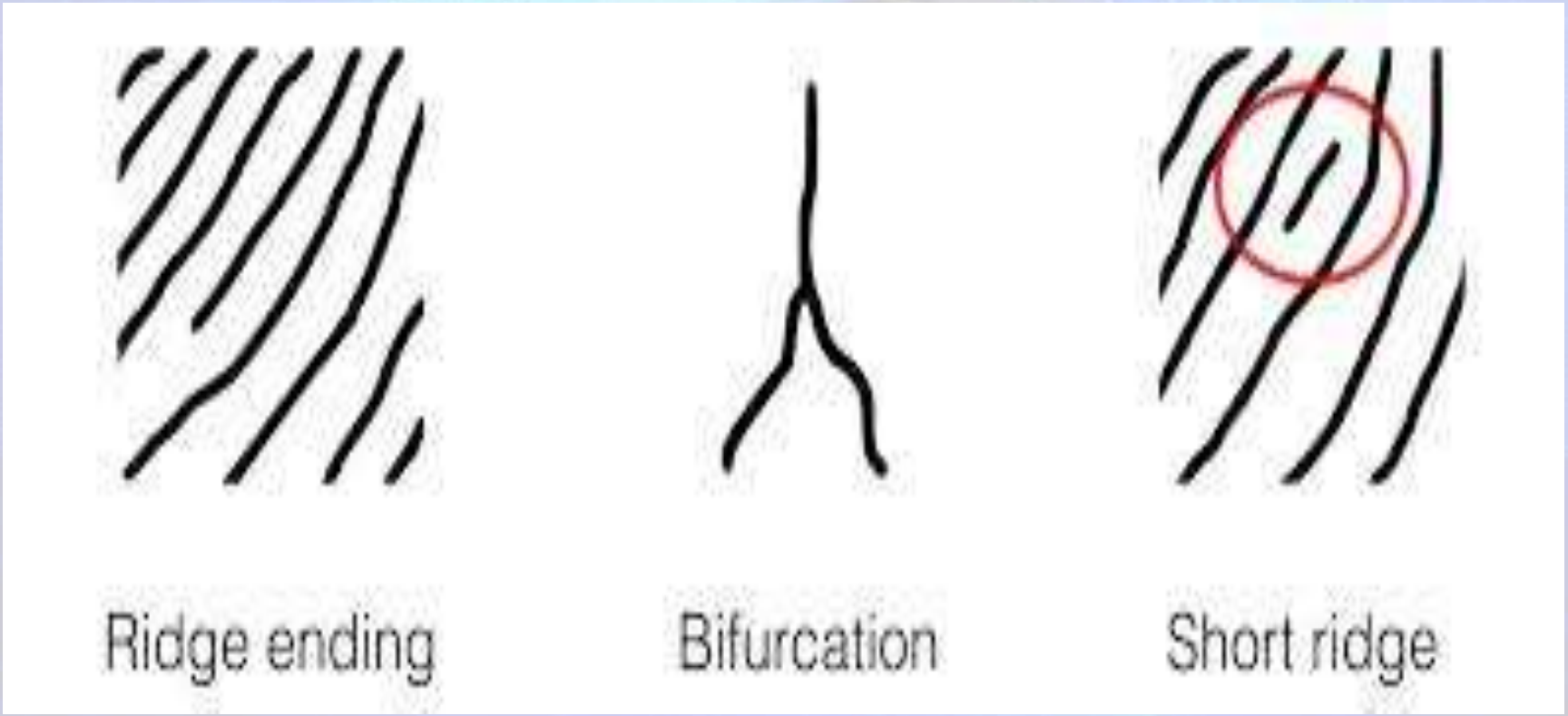
# Fingerprint

- The popular Biometric used to authenticate person is Fingerprint which is unique and permanent throughout a person's life
- Fingerprint is the pattern of ridges and valleys
- The ridges have characteristics, called minutiae, are the ridge ending and the ridge bifurcation
- Ridge ending is defined as the point where ridge ends abruptly
- Ridge bifurcation is defined as the point where a ridge forks into branch ridges

# Valleys and Ridges



# Ridge Ending and Bifurcation



# Fingerprint Recognition

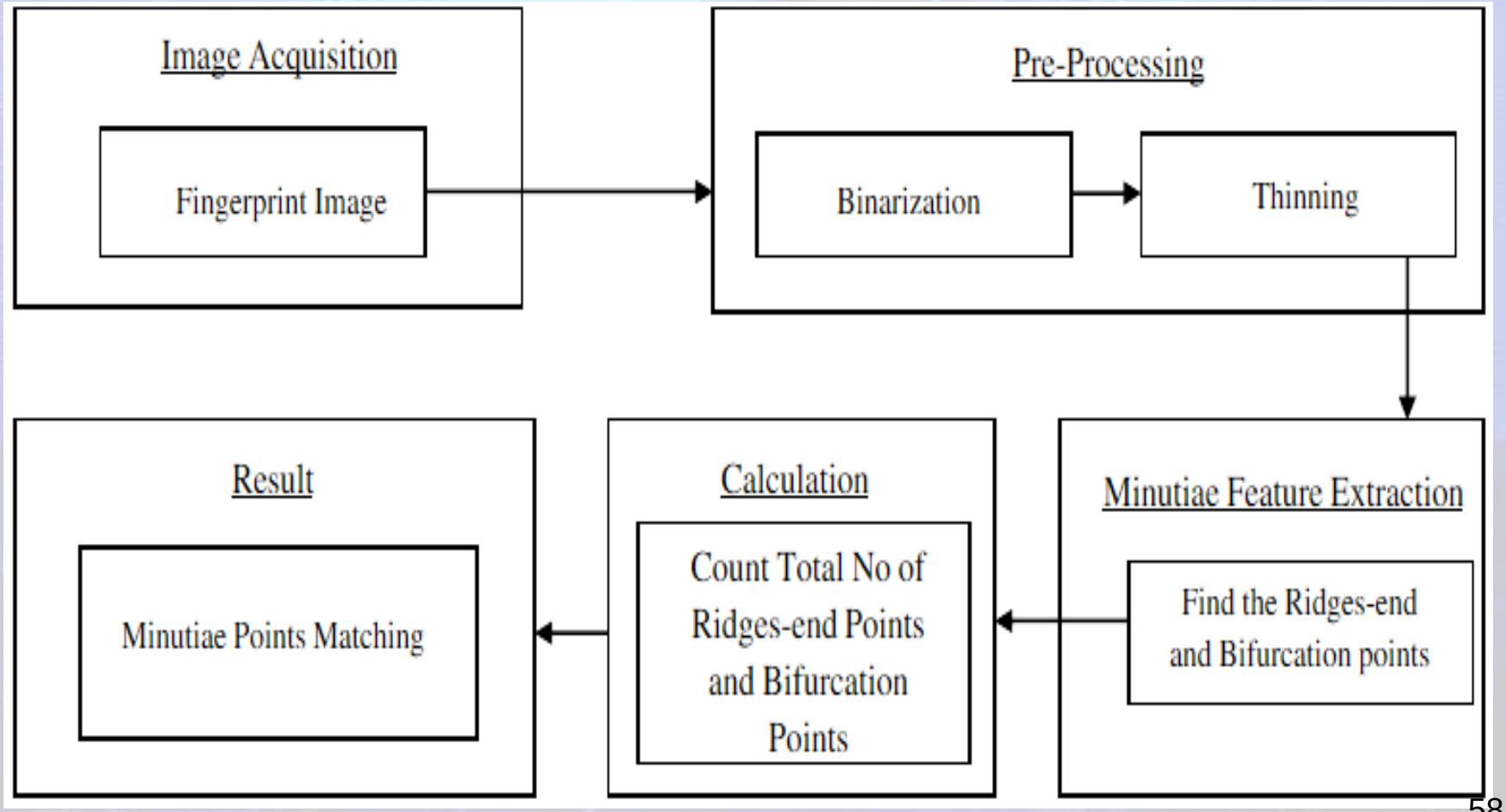
Fingerprint recognition or fingerprint authentication refers to the method of verifying a match between two human fingerprint

Fingerprint recognition techniques have the advantage to use low-cost standard capturing device

However , recognition of the fingerprint becomes a complex computer vision problem , especially when dealing with noisy and low quality images

A minutia matching is widely used for fingerprint recognition and can be classified as ridge ending and ridge bifurcation

# Fingerprint Matching using Ridge-End and Bifurcation Points



# Fingerprint Image

- The input fingerprint image is the gray scale image of a person, which has intensity values ranging from 0 to 255
- A number of methods are used to acquire fingerprints
- The inked impression method remains the most popular one
- Inkless fingerprint scanners are also present

Inked method

Inkless method

# Binarization

Binarization is used to convert gray scale image into binary image by fixing the threshold value

The pixel values above the threshold are set to ‘1’ and the pixel values below the threshold are set to ‘0’ respectively

Original Fingerprint

Binarized Fingerprint

# Thinning

The binarized image is thinned using Block Filter

To reduce the thickness of all ridge lines to a single pixel width to extract minutiae points effectively

Thinning does not change the location of minutiae points compared to original fingerprint

**Binarized Fingerprint** **Image after Thining**

# Minutiae Extraction

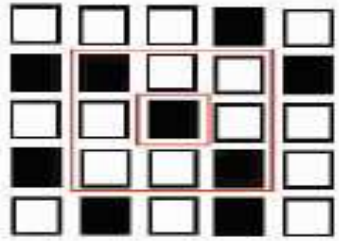
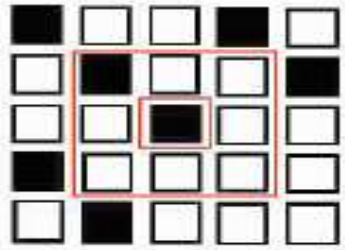
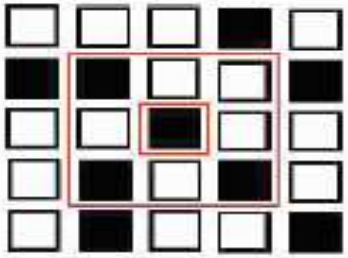
Classification of ridge-end and ridge bifurcation points is done by creating matrix

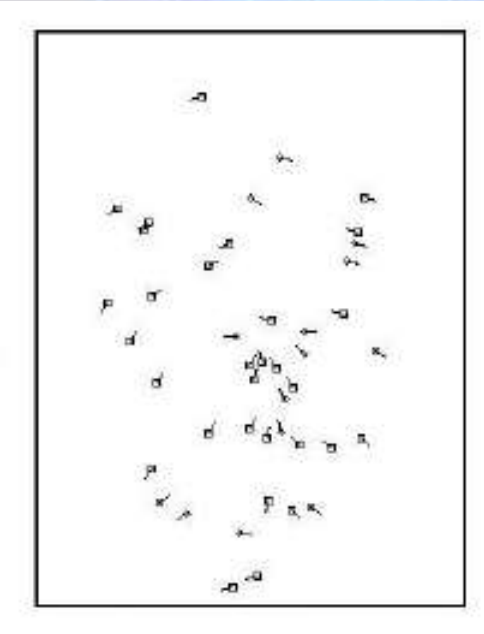
Crossing Number is used to locate the minutiae points in fingerprint image

Crossing Number is defined as half of the sum of differences between intensity values of two adjacent pixels

- If crossing Number is 1 → minutiae points are classified as Termination
- If crossing Number is 2 → minutiae points are classified as Normal ridge
- If crossing Number 3 or greater than 3 → minutiae points are classified as Bifurcation

# Crossing Number and Type of Minutiae

| | |
|---|---|
|  | Crossing Number =2.<br>Normal ridge pixel. |
|  | Crossing Number =1.<br>Termination point. |
|  | Crossing Number =3.<br>Bifurcation point. |



Gray-scale Fingerprint

Minutiae points

# Minutiae Matching

- Image Acquisition
- Computation of Points
- Location Detection of Points
- Amount and Location Matching

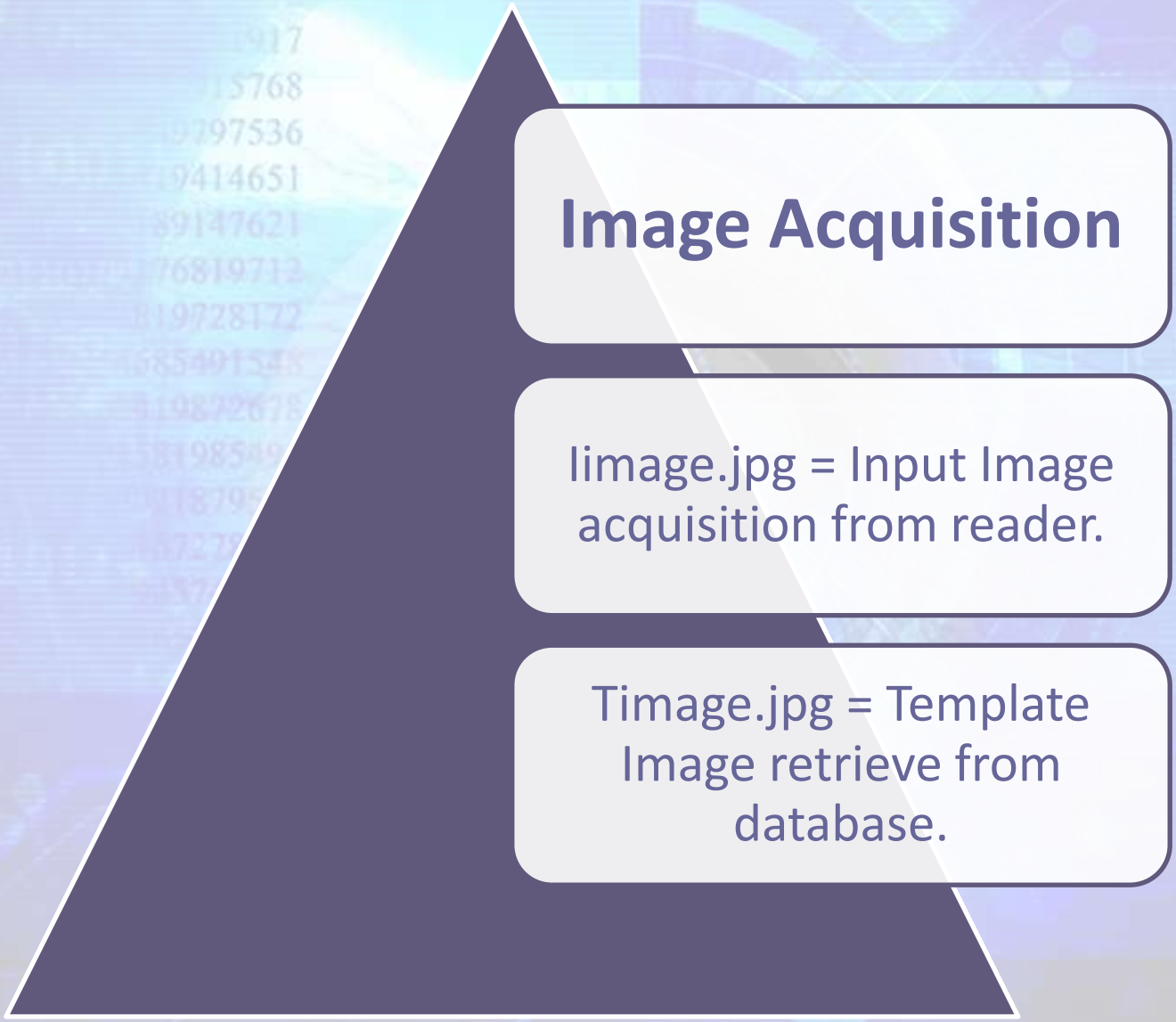


Image Acquisition
Iimage.jpg = Input Image acquisition from reader.
Timage.jpg = Template Image retrieve from database.

# Computation of Points

After the detection of minutiae points, matching algorithm require to calculate total number of available points in the fingerprint image separately

To perform this computation two counter variables are used to count both ridge-end and bifurcation points

# Minutiae Point Calculation

| Sr. No | Images | Ridges Points | Bifurcation Points |
|---|---|---|---|
| 1. | Iimage.jpg | 545 | 2858 |
| 2. | Timage.jpg | 161 | 860 |

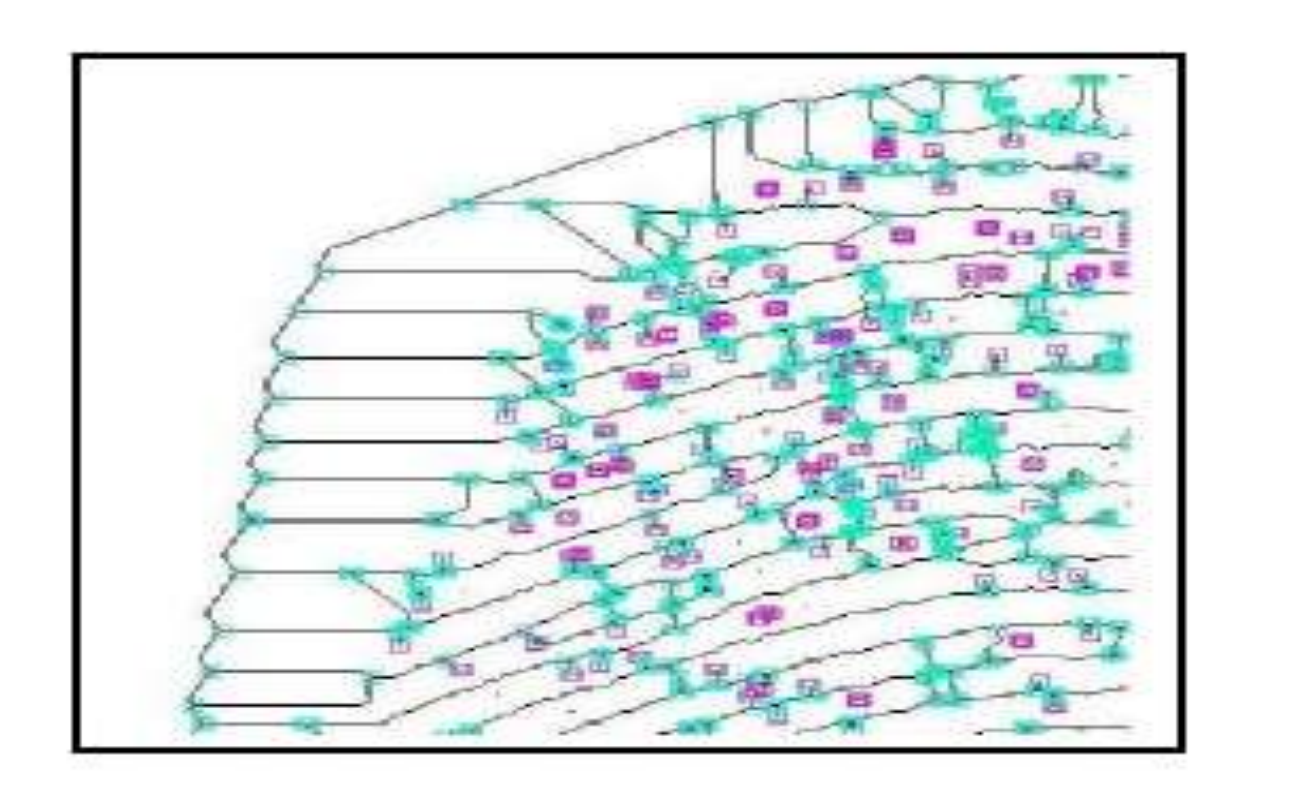
# Location Detection of Points

Each minutiae point in the fingerprint image has a specific location.

This location information of particular point is significant to store for further matching of fingerprints.

The location of every point in the digital image is given by pixel position, so that it can be taken and stored separately for both ridge-end and bifurcation points.

# Minutiae Point Extracted in Input Image



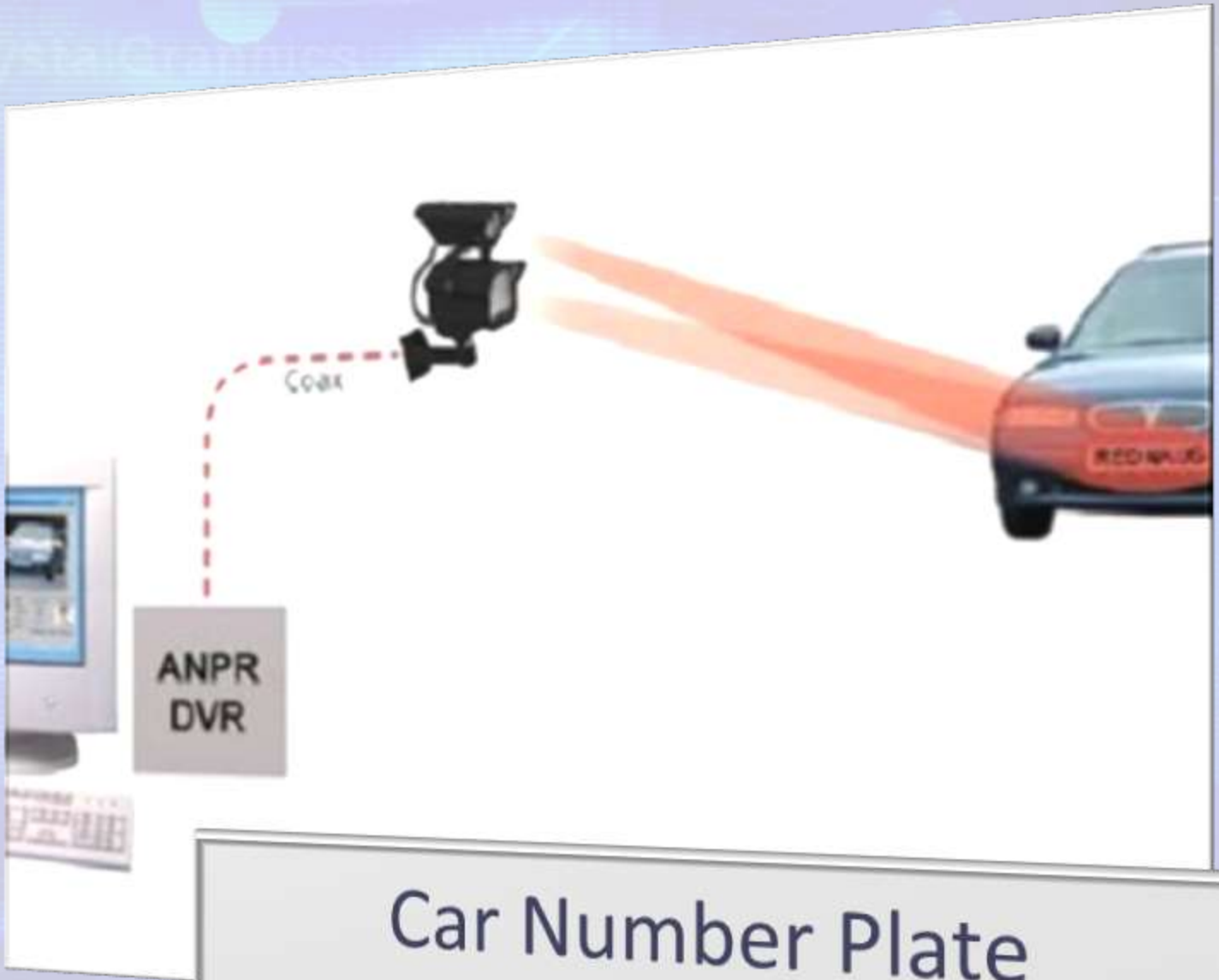
# Amount and Location Matching

In the previous steps, all the required information about points is computed and stored

Now, this is the matching step, here the algorithm compares the computed values with the stored values

This algorithm first, compares the combination of both amounts of ridge-end and bifurcation points with stored data

If the match occurs, the algorithm then compares the location of ridge points with stored location data




# Car Number Plate Recognition

# Other Names

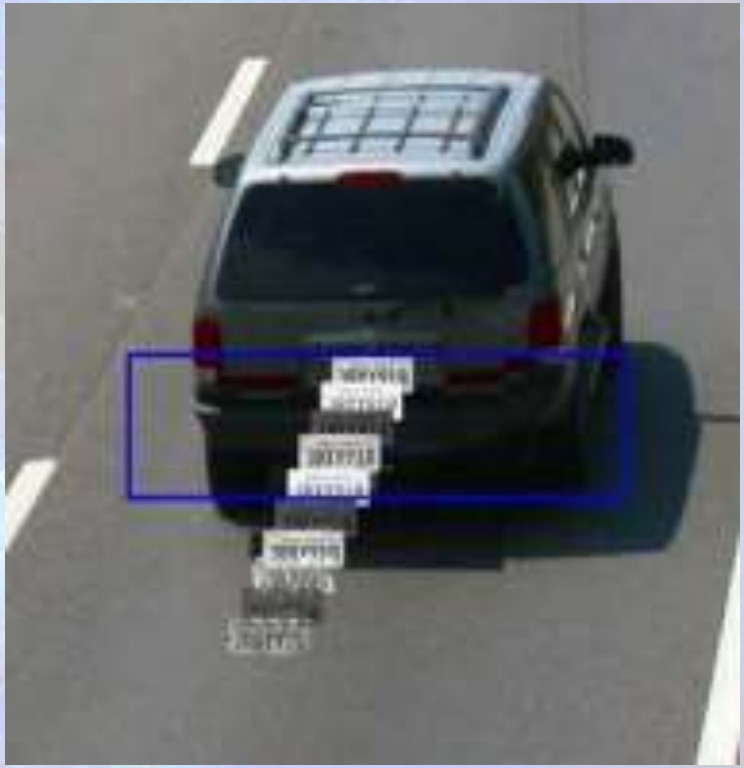
- license-plate recognition (ALPR)
- Automatic license-plate reader (ALPR)
- Automatic vehicle identification (AVI)
- Automatic Number plate recognition (ANPR)
- License-plate recognition (LPR)
- Lecture Automatique de Plaques d'Immatriculation (LAPI)
- Mobile license-plate reader (MLPR)

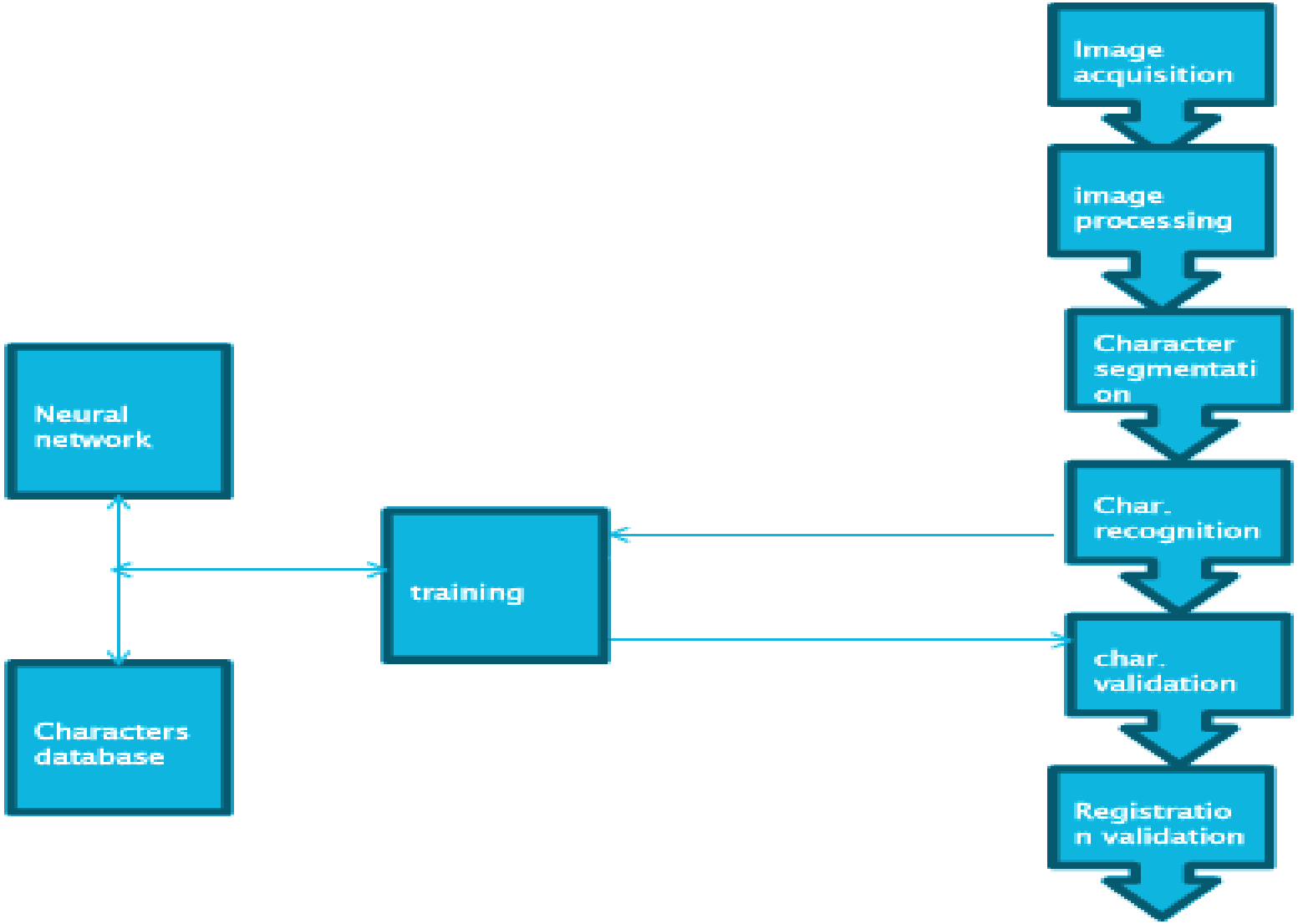
# What is an ALPR System?

Essentially it is a combination of cameras , illumination and a suite of algorithms that delivers the license plate number from a passing vehicle.

# What is an ALPR System?

- Some toll road requirements we encounter:
  – 99.9% image capture
  – 99% overall plate read accuracy
  – 99% OCR accuracy on 90% of capture plate images – number and state






The procedure is based on extraction of plate region, segmentation of plate characters and recognition of characters

## Characters Recognition

- Image binarization Image of the car p late number captured from the camera is first converted to the grascaled image
- Binary image consisting of only 1's and 0's (only black and white).
- The essential point for binarization is the thresholding . by thresholding the pixel values of 0 (black) for all of image with luminance less than threshold value and 1 (white) for all other pixels.

# Characters Segmentation

- In the segmentation of plate characters, the car number plate is segmented into its constituent parts to obtain its characters individually . Image filtering from unwanted spots and noise.

- Dilation of image to separate characters from each others.

Segmented plate number

# Separating the plate characters

It is done by finding starting and end points of characters in horizontal direction.


**Characters separated individually**

# Normalization of characters

Normalization is to refine the characters into a block containing no extra white spaces (pixels) in all the four sides of the characters.

Sometimes called contrast stretching.

Then each character should be equal in size.

Below an example of normalized character where the character fill all the 4 sides

Normalized character


LOOK
WHAT'S
NEW!

# Supermarket scanner ***recognizes objects*** without barcodes

The Object Recognition Scanner recognize items based on color and pattern

# Supermarket scanner recognizes objects without barcodes

Uses object recognition to identify foods at the supermarket checkout line.

The technology uses a camera that compares the food that is being scanned to a large, expandable database of products.

That camera filters out background "noise" in its picture, so that it only sees objects held close to its lens against a neutral black background.

The technology recognizes supermarket items at check out without requiring a bar code...making bar codes obsolete for check out purposes.

It uses proprietary pattern recognition technology and claims it can operate at high speeds.

This object recognition system requires a database that contains the information about the items in the supermarket.

This system claims to be able to make very precise identification of produce.

# Google patents new ***object recognition technology***, likely has plans to use with YouTube

# Google patents new object recognition technology, likely has plans to use with YouTube

It's known as "automatic large scale video object recognition."

It can actually recognize the difference between a variety of objects, not just human faces.

After recognizing an object it then labels it with certain tags. If you are wondering how it does this, there is a special object name repository involved.

This database would hold at least 50,000 object names, information and shapes that would allow for easy identification.

# Android Eye


Picture taken by camera:
Android Eye's guess:
"Pretty Girl"
Starbucks Now"
Android Eye's guess:
COFFEE

# Android Eye

Android Eye is an advanced Object Recognition app. Take a picture of any object, and Android Eye will tell you what it is.

Take a picture of a car... Android Eye will tell you the make and model of the car.

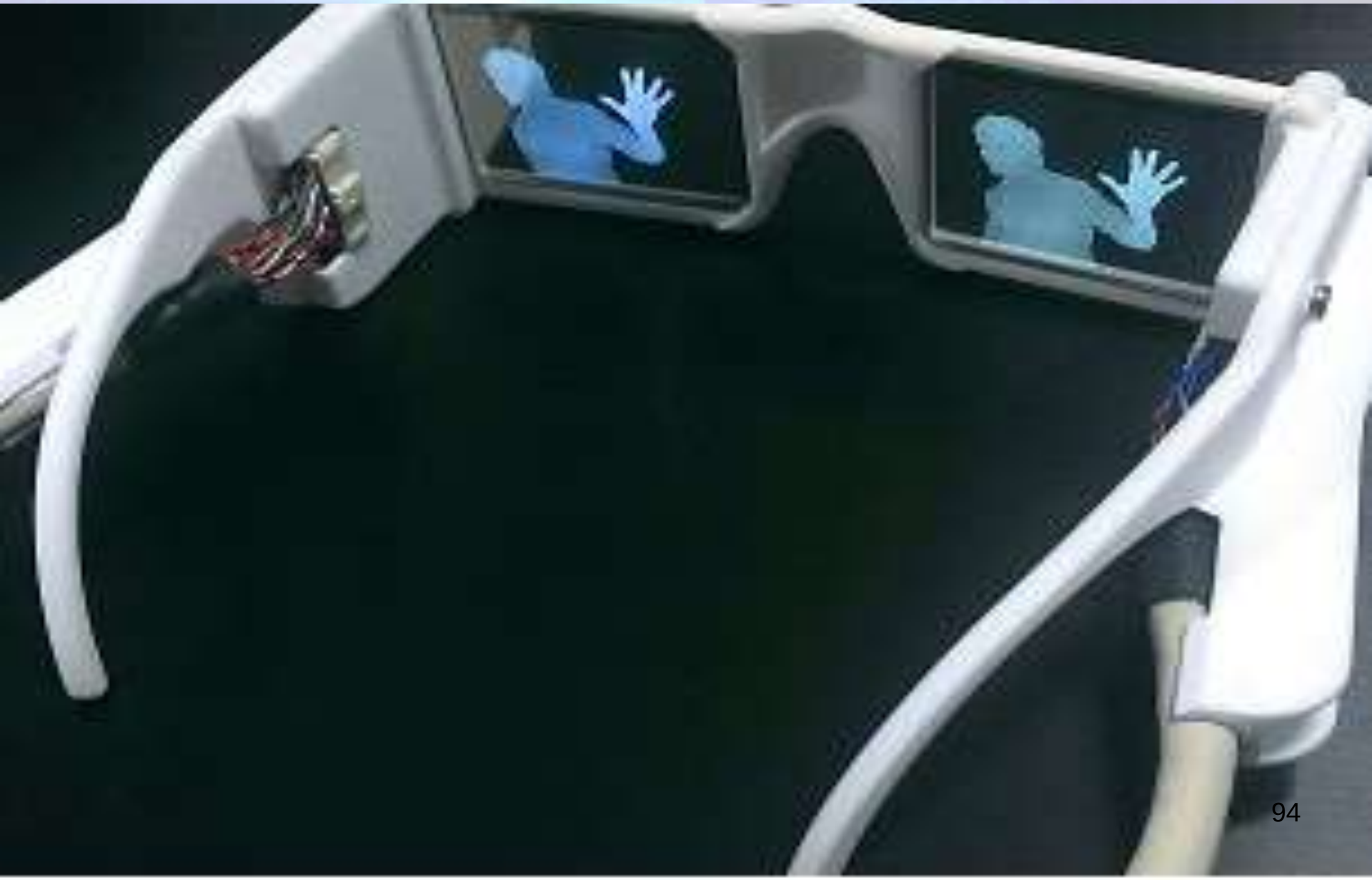
Take a picture of a foreign t-shirt label... Android Eye will tell you the brand, and where the shirt is from. Take a picture of a tree... a ball... a person... the results are endless.

This is the new version; it works very well, particularly with vehicles, products, brands, and well-known "things".

It also guesses celebrity names.

Software that does this is usually only available to government agencies and research facilities.

# A Google Glass App



# A Google Glass App Knows What You're Looking At

- An app for Google's wearable computer Glass can recognize objects in front of a person wearing the device.
- Google has shown that the camera integrated into Google Glass, the company's head-worn computer, can capture some striking video.
- They built an app that uses that camera to recognize what a person is looking at.
- The app was built at an employee hack session held by the company this month to experiment with ways to demonstrate their new image recognition service.
- The app can either work on photos taken by a person wearing Glass, or constantly grab images from the device's camera.
- Those are sent to the cloud or a nearby computer for processing by AlchemyAPI's image recognition software. The software sends back its best guess at what it sees and then Glass will display, or speak, the verdict.

# SELF-PARKING CAR

No humans here. This Passat is parking itself.

# SELF-PARKING CAR

- The autonomous valet parking is just one aspect of a greater effort to create fully autonomous vehicles.
- As object recognition software improves, and decision making capabilities are refined, cars will become capable of increasingly complex driving routines.
- The vehicle uses a camera placed in front of the rear view mirror and a front radar system
- Small lasers sensors were added along the periphery
- The VAIL/Stanford car can park itself, but only if provided with a map of the lot beforehand.
- The robotic car is unable to avoid obstacles and will not stop if encounters one.

# References


- http://arxiv.org/ftp/arxiv/papers/1001/1001.4186.pdf
- http://www4.comp.polyu.edu.hk/~csajaykr/myhome/teaching/biometrics/final_report.pdf
- http://www.google.com.lb/ Fingerprint Analysis and Representation.ppt
- http://www.academia.edu/2537762/Fingerprint_Matching_using_Ridge-End_and_Bifurcation_Points
- http://ai.pku.edu.cn/aiwebsite/research.files/collected%20papers%20-%20fingerprint/Minutiae%20detection%20algorithm%20for%20fingerprint%20recognition.pdf
- http://en.wikipedia.org/wiki/Computer_vision


thank
you!