

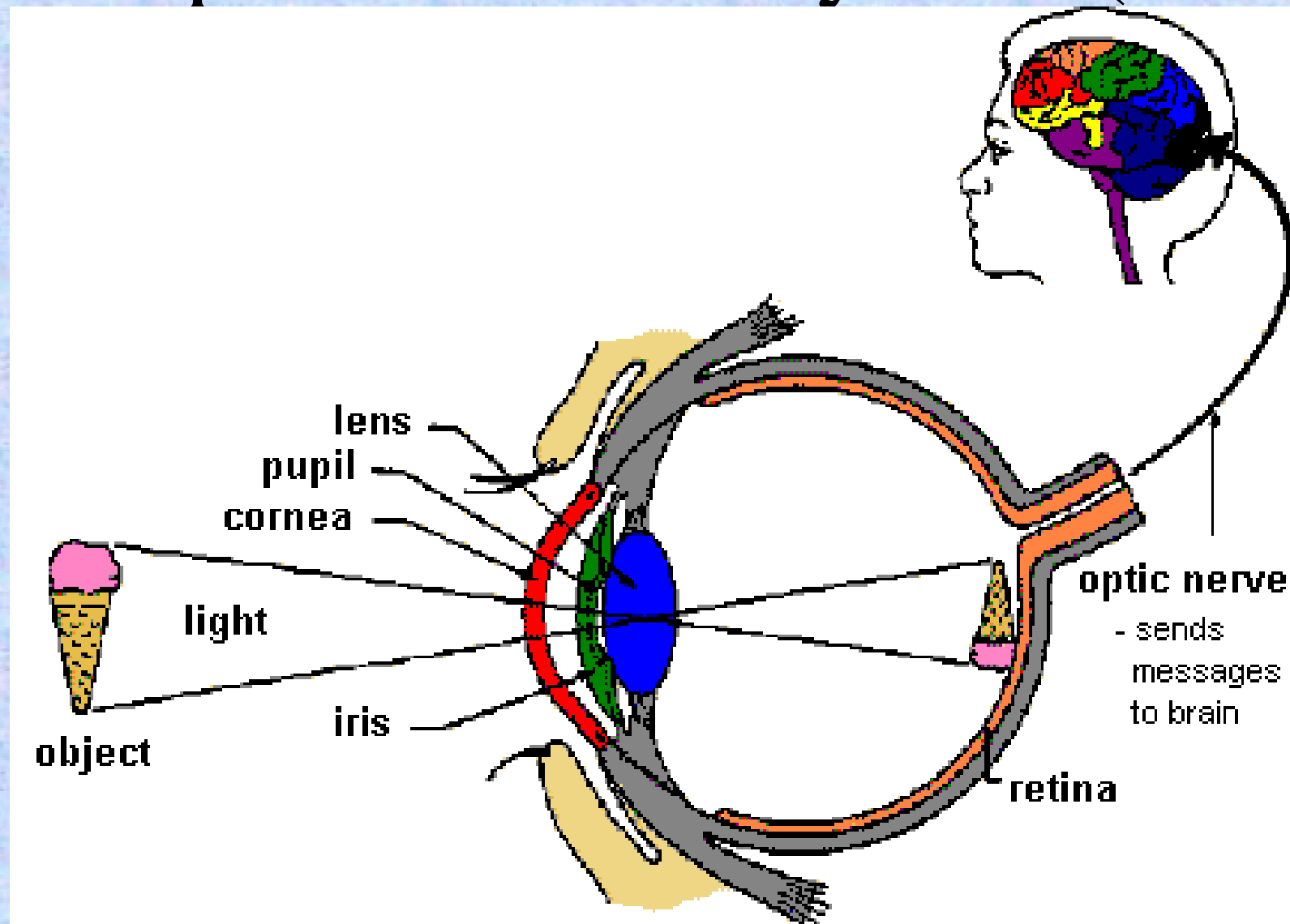
COMPUTER VISION

Dr. Sukhendu Das
Deptt. of Computer Science and Engg.,
IIT Madras, Chennai – 600036.

Email: sdas@iitm.ac.in
URL: [//www.cs.iitm.ernet.in/~sdas](http://www.cs.iitm.ernet.in/~sdas)

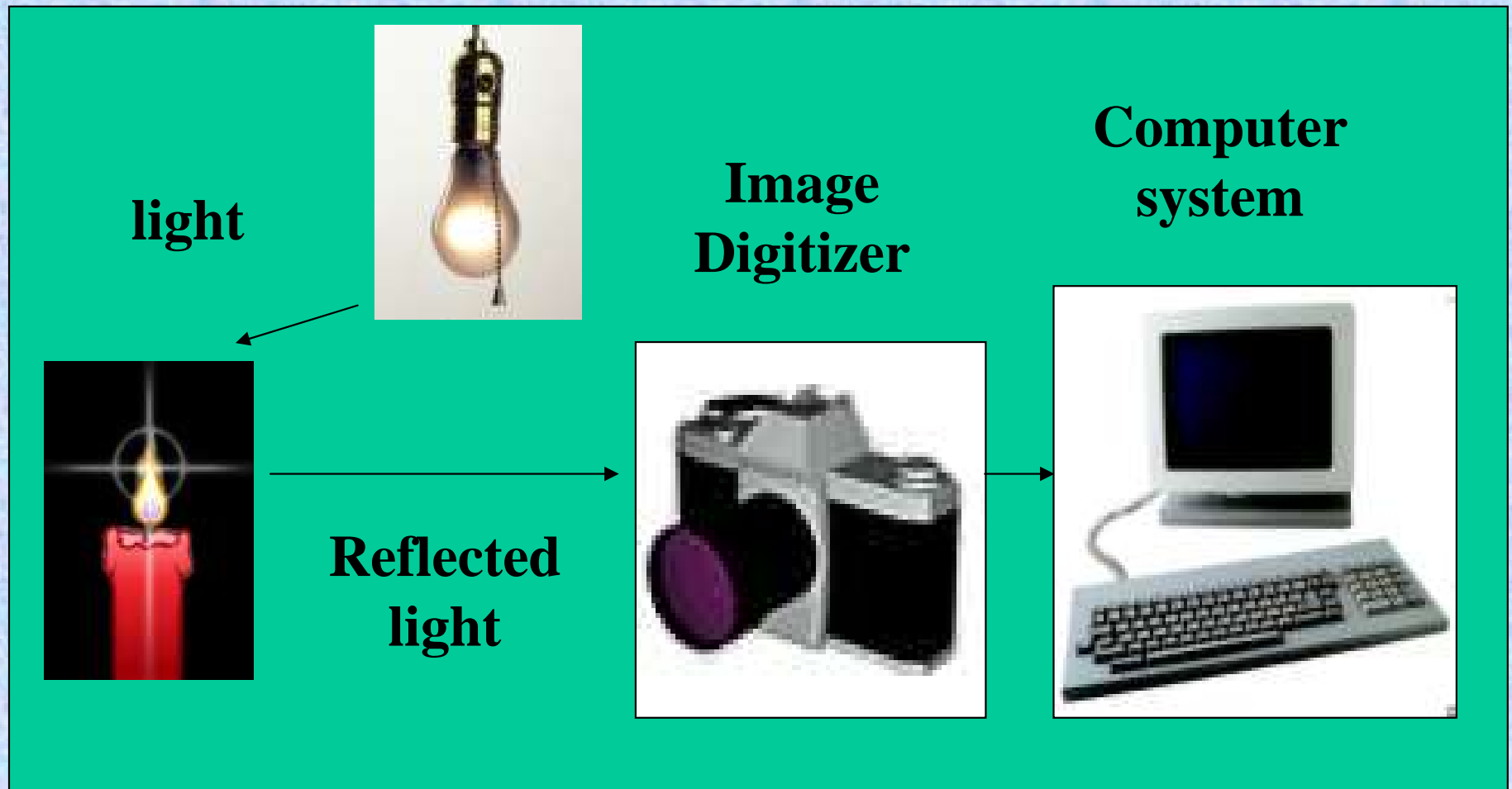
INTRODUCTION

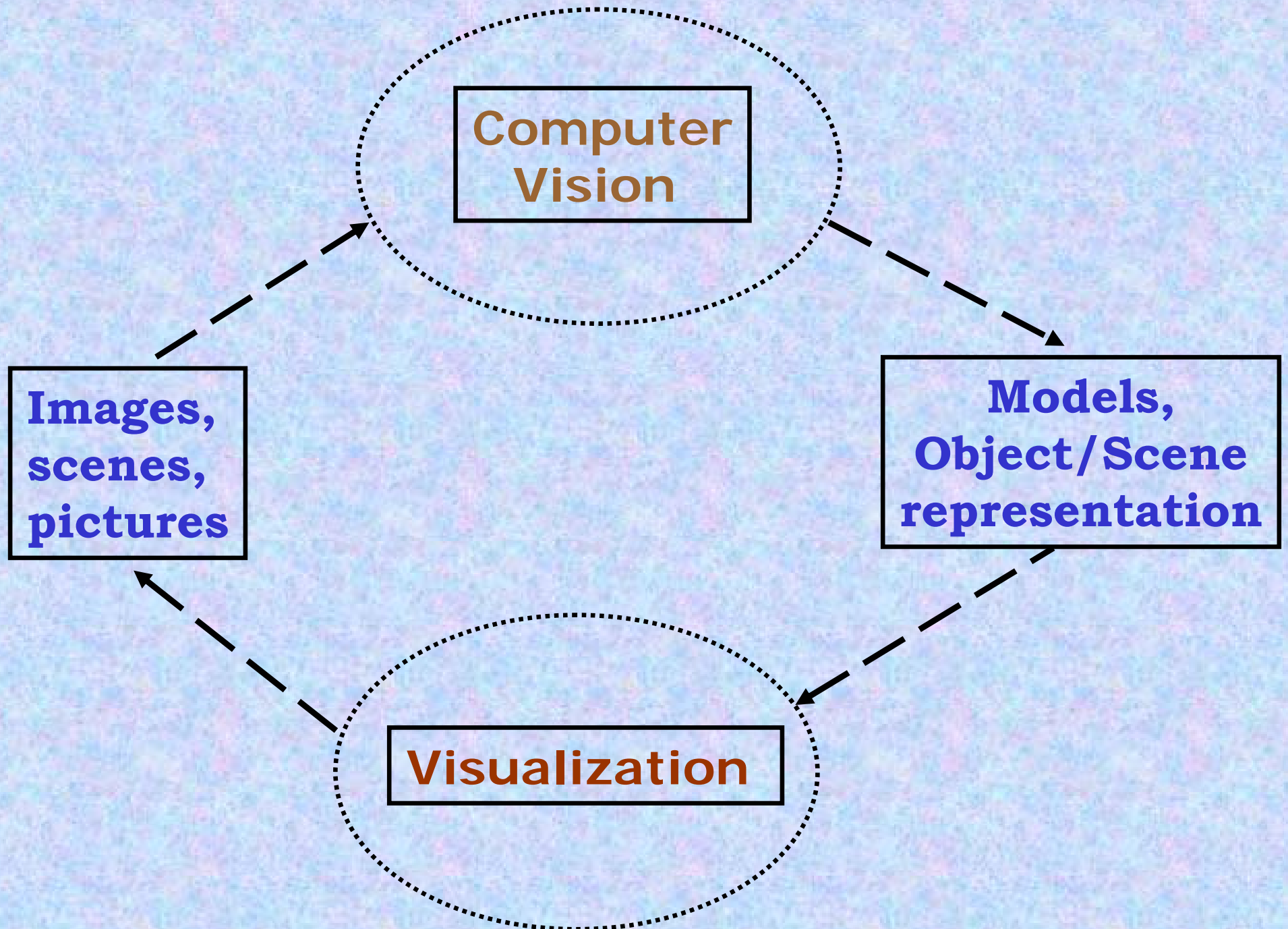
Human Vision System (HVS) Vs. Computer Vision System (CVS)



The Optics of the eye

A computer Vision System (CVS)



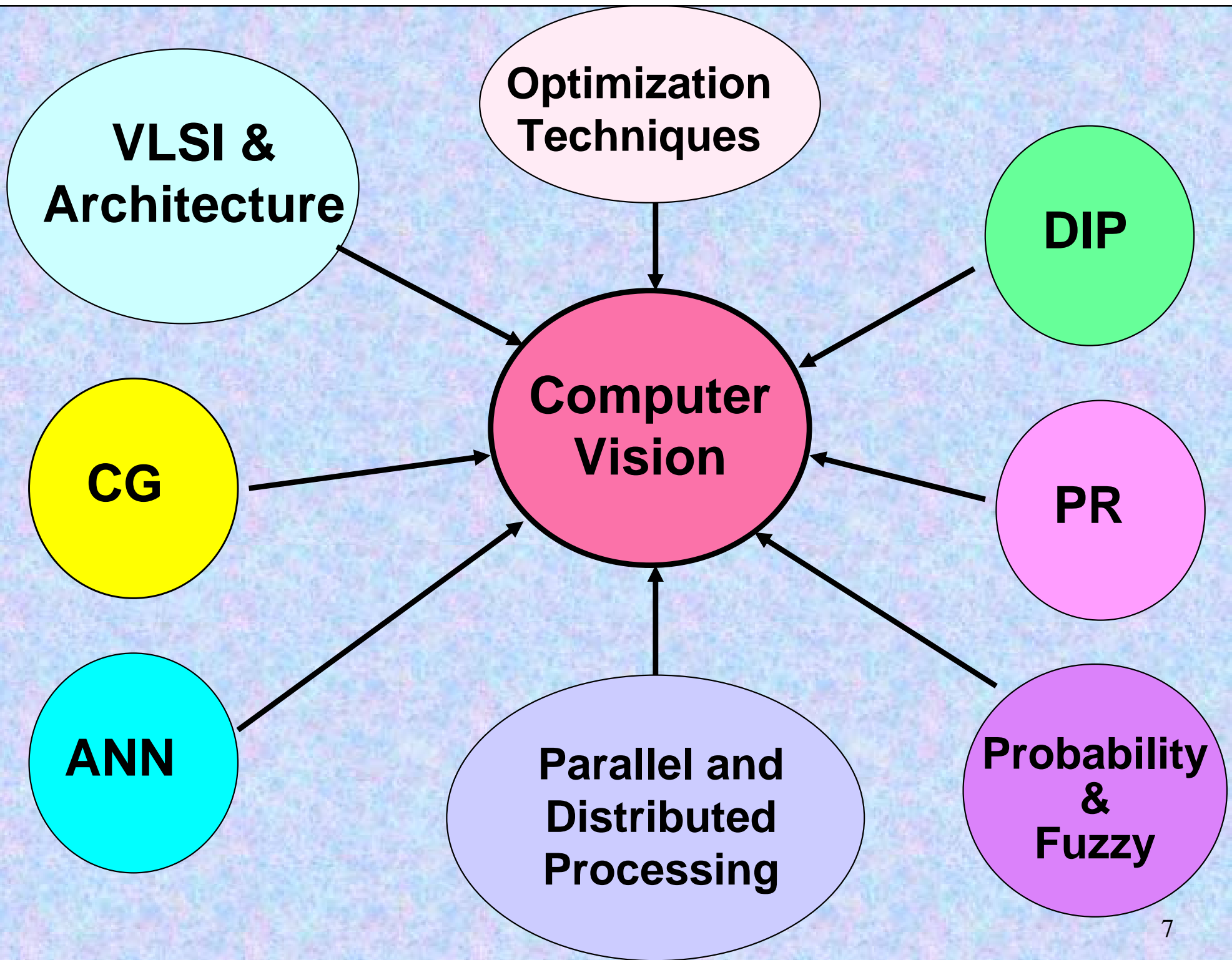


Computer Vision is an area of work, which is a combination of concepts, techniques and ideas from Digital Image Processing, Pattern Recognition, Artificial Intelligence and Computer Graphics.

Majority of the tasks in the fields of Digital Image Processing or Computer Vision deals with the process of **understanding** or deriving the scene information or description, from the input scene (digital image/s). The methods used to solve a problem in digital image processing depends on the **application domain** and nature of data being analyzed.

Analysis of two-dimensional pictures are generally not applicable of processing three-dimensional scenes, and vice-versa. The choice of processing, techniques and methods and '**features**' to be used for a particular application is made after some amount of trial and error, and hence experience in handling images is crucial in most of these cases.

For example, analysis of remote sensed or satellite imagery involves techniques based on classification or analysis of texture imagery. These techniques are not useful for analyzing optical images of indoor or outdoor scenes.



Digital Image processing is in many cases concerned with taking one array of pixels as input and producing another array of pixels as output which in some way represents an improvement to the original array.

Purpose:

1. Improvement of Pictorial Information

- improve the contrast of the image,
- remove noise,
- remove blurring caused by movement of the camera during image acquisition,
- it may correct for geometrical distortions caused by the lens.

2. Automatic Machine perception (termed Computer Vision, Pattern Recognition or Visual Perception) for intelligent interpretation of scenes or pictures.

Elements of a Digital Image Processing System

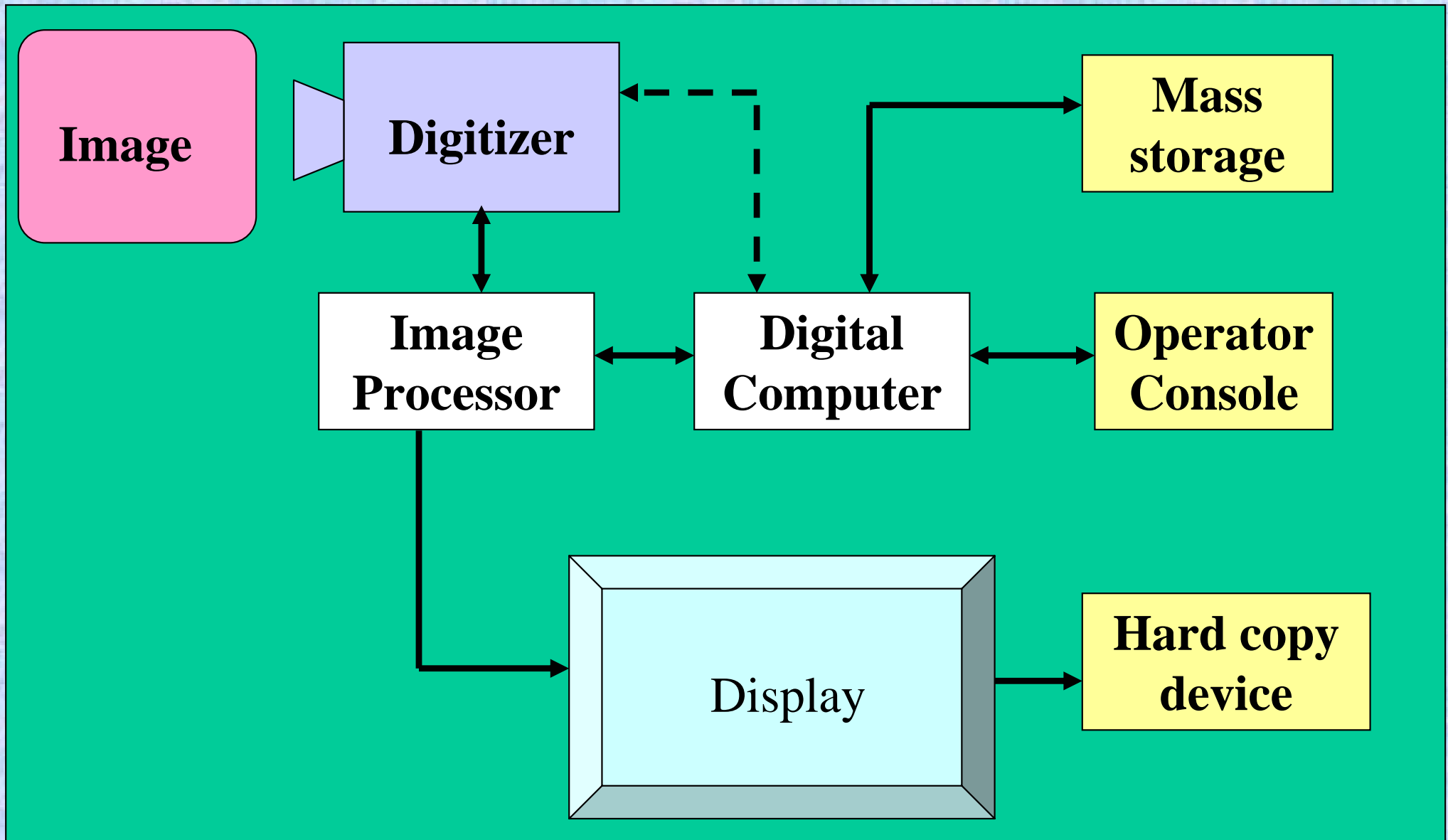
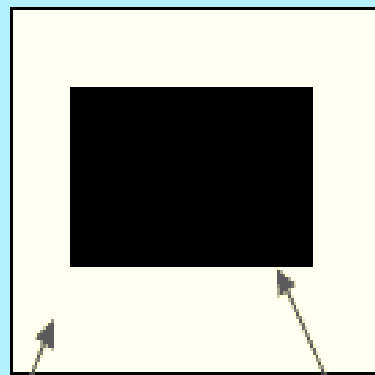


Image processors: Consists of set of hardware modules that perform 4 basic functions:

- Image acquisition: frame grabber**
 - Storage: frame buffer**
 - Low-level processing: specialized hardware device designed to perform Arithmetic Logic operations on pixels in parallel**
 - Display: read from image memory (frame buffer) and convert to analog video signal**
- Digitizers: Converts image into numerical representation suitable for input to a digital computer**
 - Digital Computers: Interfaced with the image processor to provide versatility and ease of programming.**
 - Storage Devices: For bulk storage. e.g:- Magnetic disks, magnetic tapes, optical disks**
 - Display and Recording devices : Monochrome and Color Television monitors, CRT, Laser printers, heat-sensitive paper devices, and ink spray systems.**

Image acquisition using a CCD camera

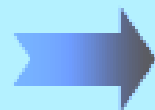
ORIGINAL SCENE



Light Areas
(high intensity)

Dark Areas
(low intensity)

*Image capture -
'sampling' using
digital camera
or scanner*



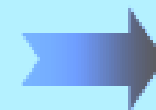
CCD Response

✓	✓	✓	✓	✓	✓
✓	x	x	x	x	✓
✓	x	x	x	x	✓
✓	✓	✓	✓	✓	✓

Electrical response
proportional to light
exposure

No response in
unexposed
areas

*Analogue to digital
conversion followed
by conversion to
image pixel values
for display*



Digital Image

255	255	255	255	255	255
255	0	0	0	0	255
255	0	0	0	0	255
255	255	255	255	255	255

Pixel
representations of
white and black

A digital Image

Image is an array of integers: $f(x,y) \in \{0,1,\dots,I_{\max}-1\}$,
where, $x,y \in \{0,1,\dots,N-1\}$

- **N is the resolution of the image and I_{\max} is the level of discretized brightness value**
- **Larger the value of N, more is the clarity of the picture (larger resolution), but more data to be analyzed in the image**
- **If the image is a gray-level (8-bit per pixel - termed raw, gray) image, then it requires N^2 Bytes for storage**
- **If the image is color - RGB, each pixel requires 3 Bytes of storage space.**

Image Size (resolution)	Storage space required	
	Raw - Gray	Color (RGB)
64*64	4K	12K
256*256	64K	192K
512*512	256K	768K

A digital image is a two-dimensional (3-D image is called range data) array of intensity values, $f(x, y)$, which represents 2-D intensity function discretized both in spatial coordinates (**spatial sampling**) and brightness (**quantization**) values.

The elements of such an array are called **pixels** (picture elements).

The storage requirement for an image depends on the **spatial resolution** and number of bits necessary for **pixel quantization**.

The processing of an image depends on the application domain and the methodology used to solve a problem. There exists four broad categories of tasks in digital image processing:

- | | |
|-----------------------|--------------------|
| (i) Compression, | (ii) Segmentation, |
| (iii) Recognition and | (iv) motion. |

Segmentation deals with the process of fragmenting the image into homogeneous meaningful parts, regions or sub-images. Segmentation is generally based on the analysis of the histogram of images using gray level values as features. Other features used are edges or lines, colors and textures.

Recognition deals with identification or classification of objects in an image for the purpose of interpretation or identification. Recognition is based on models, which represent an object. A system is trained (using HMM, GMM, ANN etc.) to learn or store the models, based on training samples. The test data is then matched with all such models to identify the object with a certain measure of confidence.

Compression involves methodologies for efficient storage and retrieval of image data, which occupies large disk space. Typical methods are, JPEG-based, Wavelet based, Huffman Coding, Run length coding etc. for still images and MPEG-I, II, IV & VII for digital video or sequence of frames.

Motion analysis (or dynamic scene analysis) involves techniques for the purpose of tracking and estimation of the path of movement of object/s from a sequence of frames (digital video). Methods for dynamic scene analysis are based on (i) tracking, (ii) obtaining correspondence between frames and then (iii) estimating the motion parameters and (iv) structure of moving objects. Typical methods for analysis are based on optical flow, iterative Kalman filter and Newton/Euler's equations of dynamics.

There are generally three main categories of tasks involved in a complete computer vision system. They are:

- *Low level processing:* Involves image processing tasks in which the quality of the image is improved for the benefit of human observers and higher level routines to perform better.
- *Intermediate level processing:* Involves the processes of feature extraction and pattern detection tasks. The algorithms used here are chosen and tuned in a manner as may be required to assist the final tasks of high level vision.
- *High level vision:* Involves autonomous interpretation of scenes for pattern classification, recognition and identification of objects in the scenes as well as any other information required for human understanding.

A top down approach, rather than a bottom-up approach is used in the design of these systems in many applications.

Different fields of applications include:

- Character Recognition,
- Document processing,
- Commercial (signature & seal verification) application,
- Biometry and Forensic (authentication: recognition and verification of persons using face, palm & fingerprint),
- Pose and gesture identification,
- Automatic inspection of industrial products,
- Industrial process monitoring,
- Biomedical Engg. (Diagnosis and surgery),
- Military surveillance and target identification,
- Navigation and mobility (for robots and unmanned vehicles - land, air and underwater),
- Remote sensing (using satellite imagery),
- GIS
- Safety and security (night vision),
- Traffic monitoring,
- Sports (training and incident analysis)
- VLDB (organization and retrieval)
- Entertainment and virtual reality.

The various sub-categories of work in these related fields are:

- *image enhancement,*
- *image restoration and filtering,*
- *representation and description,*
- *feature extraction,*
- *image segmentation,*
- *image matching,*
- *color image processing,*
- *image synthesis,*
- *image representation,*
- *image reconstruction*
- *range data processing,*
- *stereo image processing*
- *computational geometry,*
- *image morphology,*
- *artificial neural networks,*
- *Neuro-fuzzy techniques,*
- *computational geometry,*
- *parallel architectures & algorithms.*

References

1. **“Digital Image Processing”; R. C. Gonzalez and R. E. Woods; Addison Wesley; 1992+.**
2. **“Computer Vision: A Modern Approach”; D. A. Forsyth and J. Ponce; Pearson Education; 2003.**
3. **“Fundamentals of Digital Image Processing”; Anil K Jain; Prentice Hall of India; 1995+.**
4. **“3-D Computer Vision”; Y. Shirai; Springer-Verlag, 1984.**
5. **“Digital Image Processing and Computer Vision”; Robert J. Schalkoff; John Wiley and Sons; 1989+.**
6. **“ Pattern Recognition: Statistical. Structural and Neural Approaches”; Robert J. Schalkoff; John Wiley and Sons; 1992+.**

References (Contd..)

Journals:

- IEEE-T-PAMI (Transactions on Pattern Analysis and Machine Intelligence)
- IEEE-T-IP (Transactions on Image processing)
- PR (Pattern Recognition)
- PRL (Pattern Recognition Letters)
- CVGIP (Computer Vision, Graphics & Image Processing)
- IJCV (International Journal of Computer Vision)

Online links

1. CV online: <http://homepages.inf.ed.ac.uk/rbf/CVonline>
2. Computer Vision Homepage:
<http://www-2.cs.cmu.edu/afs/cs/project/cil/ftp/html/vision.html>