Computer vision notes part 1

Computer vision (Islamic University of Science and Technology)

UNIT I

What is 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.
Computer vision is used in industries ranging from energy and utilities to manufacturing and automotive – and the market is continuing to grow.

What is digital image processing?

digital image processing is the processing of images that are digitally stored using a computer. In other words, we can say that "***Digital Image Processing refers to the processing of digital images by means of a digital computer.***"

IMAGE:

An image can be defined as a function *f(x,y)*. Here

- x and y are spatial (plane) coordinates.
- the amplitude of function *f* at any point *(x,y)* is called intensity or grey level of an image at that point.

Now ***an image is said to be a digital image when x,y and amplitude values of function f(x,y) are finite and discrete in nature***.

A digital image is composed of a finite number of elements (known as pixels), each of which has

- a particular value
- a particular location

A pixel is the smallest element in an image.

Analog Image Processing

The analog image processing is applied on analog signals and it processes only two-dimensional signals

. The images are manipulated by electrical signals. In analog image processing, analog signals can be periodic or non-periodic

Examples of analog images are television images, photographs, paintings, and medical images etc.

Digital Image Processing

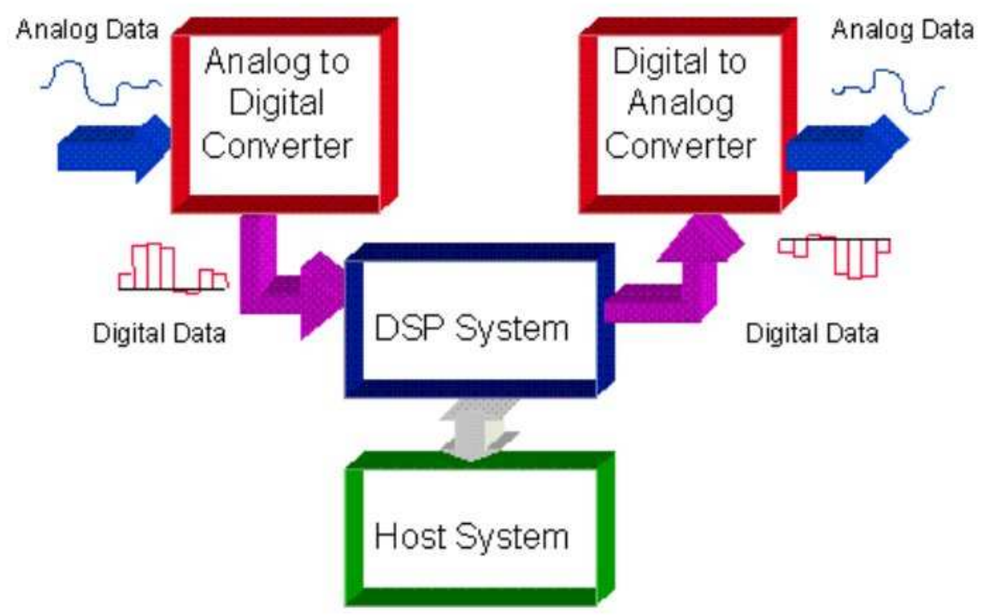
A digital image processing is applied to digital images (a matrix of small pixels and elements). For manipulating the images, there is a number of software and algorithms that are applied to perform changes. Digital image processing is one of the fastest growing industry which affects everyone's life.

Examples of digital images are color processing, image recognition, video processing, etc.

Conversion of the analog signal to a digital signal by digital image processing

Digital signal processing is all about processing analog signal or real-world signals which humans interact, for example, speech.

DSP system converts digital signal to analog signal or vice-versa by the use of converters.



The digital image processing is a special type of processor which is used in every electronic device whether it be CD, mobile phones, battlefields, satellites, medical, and voice detection machines etc.

Analog Image Processing vs. Digital Image Processing

There are following differences between Analog Images Processing and Digital Image Processing:

Analog Image Processing	Digital Image Processing
The analog image processing is applied on analog signals and it processes only two-dimensional signals.	The digital image processing is applied to digital signals that work on analyzing and manipulating the images.
Analog signal is time-varying signals so the images formed under analog image processing get varied.	It improves the digital quality of the image and intensity distribution is perfect in it.
Analog image processing is a slower and costlier process.	Digital image processing is a cheaper and fast image storage and retrieval process.
Analog signal is a real-world but not good quality of images.	It uses good image compression techniques that reduce the amount of data required and produce good quality of images
It is generally continuous and not broken into tiny components.	It uses an image segmentation technique which is used to detect discontinuity which occurs due to a broken connection path.

Three levels of processing in digital image processing:

In general, there are three levels of processing or three types of processes in digital image processing namely: low, mid and high-level processes.

- **Low-level processing** involves primitive operation such as image preprocessing to reduce noise, contrast enhancement, image sharpening, etc. In the low-level process, both input and output are images.
- **Mid-level processing** involves tasks such as image segmentation, description of images, object recognition, etc. In the mid-level process, inputs are generally images but its outputs are generally image attributes.
- **High-level processing** involves "making sense" from a group of recognized objects. This process is normally associated with computer vision.

Applications of digital image processing:

1. Automatic character recognition.
2. Military recognition.
3. Automatic processing of fingerprints.
4. Screening of X-ray and blood samples.
5. It is used in the study of pollution particles.
6. It is used to restore blurred pictures of rare artifacts (application in archaeology).
7. Industrial machine vision for products assembly and inspection.
8. Machine processing of aerial and satellite imagery for weather prediction.
9. In astronomy, biology, nuclear medicine, law enforcement, defence and industry, etc.

Advantages of digital image: *f*

-Fast Procession
-cost effective
-Efficient storage
-Transmission
-Scope for versatile
-Image manipulation

Disadvantages of digital image: *f*

-High meomory for good quality image.

Extra

Scope of image processing

While imaging was associated with security functions and surveillance missions, the term has grown to represent something larger in recent years. Thanks to advancements in science and technology, image processing is now an integral part of AI systems.

Various new types of processing systems, which have come up recently help with chemical, thermal and molecular imaging. Furthermore, the use of such systems has led to tremendous growth in the field of space exploration. Most new satellites make use of different sensors to obtain useful information from outer space. Satellite imaging and military applications are regarded as future trends in the field of image processing. Furthermore, advances in broadband devices and mobile technology will help in the improvement of image processing systems in hand-held devices. To put things in perspective, the future of image processing looks bright and solid

Fundamental Image Processing Steps

Image Acquisition

Image acquisition is the first step in image processing. This step is also known as preprocessing in image processing. It involves retrieving the image from a source, usually a hardware-based source.

Image Enhancement

Image enhancement is the process of bringing out and highlighting certain features of interest in an image that has been obscured. This can involve changing the brightness, contrast, etc.

Image Restoration

Image restoration is the process of improving the appearance of an image. However, unlike image enhancement, image restoration is done using certain mathematical or probabilistic models.

Color Image Processing

Color image processing includes a number of color modeling techniques in a digital domain. This step has gained prominence due to the significant use of digital images over the internet.

Wavelets and Multiresolution Processing

Wavelets are used to represent images in various degrees of resolution. The images are subdivided into wavelets or smaller regions for data compression and for pyramidal representation.

Compression

Compression is a process used to reduce the storage required to save an image or the bandwidth required to transmit it. This is done particularly when the image is for use on the Internet.

Morphological Processing

Morphological processing is a set of processing operations for morphing images based on their shapes.

Segmentation

Segmentation is one of the most difficult steps of image processing. It involves partitioning an image into its constituent parts or objects.

Representation and Description

After an image is segmented into regions in the segmentation process, each region is represented and described in a form suitable for further computer processing. Representation deals with the image's characteristics and regional properties. Description deals with extracting quantitative information that helps differentiate one class of objects from the other.

Recognition

Recognition assigns a label to an object based on its description.

extra

Fundamentals of Image Formation

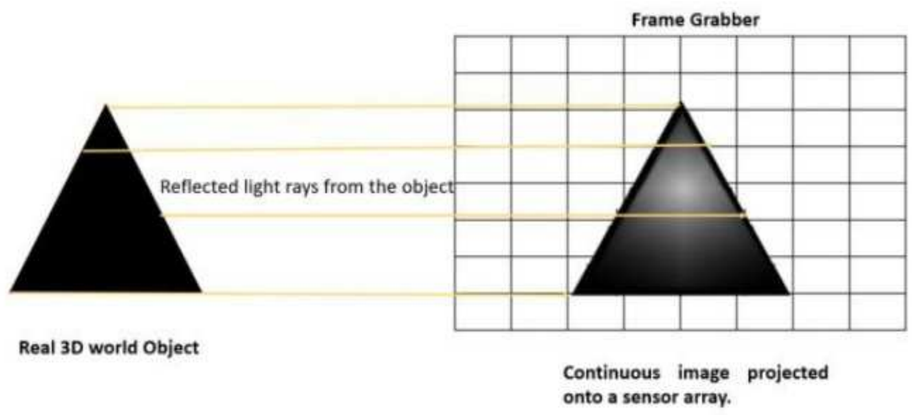
Image formation is an analog to digital conversion of an image with the help of 2D Sampling and Quantization techniques that is done by the capturing devices like cameras. In general, we see a 2D view of the 3D world.

In the same way, the formation of the analog image took place. It is basically a conversion of the 3D world that is our analog image to a 2D world that is our Digital image.

Generally, a frame grabber or a digitizer is used for sampling and quantizing the analog signals.

Imaging:

The mapping of a 3D world object into a 2D digital image plane is called *imaging*. In order to do so, each point on the 3D object must correspond to the image plane. We all know that light reflects from every object that we see thus enabling us to capture all those light-reflecting points in our image plane.
Various factors determine the quality of the image like spatial factors or the lens of the capturing device.



Color and Pixelation:

In digital imaging, a frame grabber is placed at the image plane which is like a sensor. It aims to focus the light on it and the continuous image is pixelated via the reflected light by the 3D object. The light that is focused on the sensor generates an electronic signal.

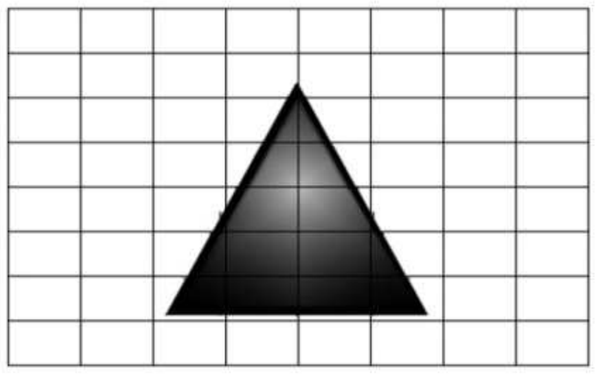
Each pixel that is formed may be colored or grey depending on the intensity of the sampling and quantization of the light that is reflected and the electronic signal that is generated via them.

All these pixels form a digital image. The density of these pixels determines the image quality. The more the density the more the clear and high-resolution image we will get.

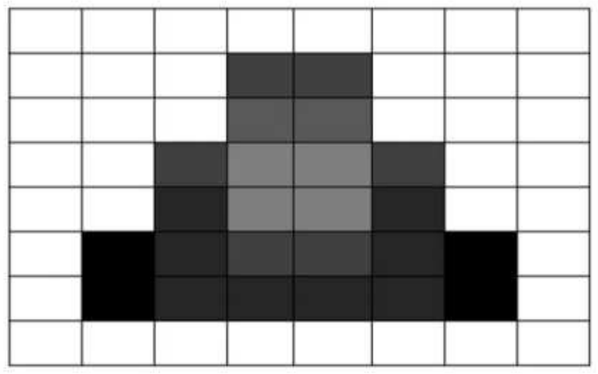
Forming a Digital Image:

In order to form or create an image that is digital in nature, we need to have a continuous conversion of data into a digital form. Thus, we require two main steps to do so:

- **Sampling (2D):** Sampling is a spatial resolution of the digital image. And the rate of sampling determines the quality of the digitized image. The magnitude of the sampled image is determined as a value in image processing. It is related to the coordinates values of the image.
- **Quantization:** Quantization is the number of grey levels in the digital image. The transition of the continuous values from the image function to its digital equivalent is called quantization. It is related to the intensity values of the image.
- The normal human being acquires a high level of quantization levels to get the fine shading details of the image. The more quantization levels will result in the more clear image.



Continuous image projected onto a sensor array.



Result of image sampling and quantization.

Types of Images

There are three types of images. They are as following:

1. Binary Images

It is the simplest type of image. It takes only two values i.e, Black and White or 0 and 1. The binary image consists of a 1-bit image and it takes only 1 binary digit to represent a pixel. Binary images are mostly used for general shape or outline

For Example: Optical Character Recognition (OCR)

Binary images are generated using threshold operation. When a pixel is above the threshold value, then it is turned white('1') and which are below the threshold value then they are turned black('0')

2. Gray-scale images

Grayscale images are monochrome images, Means they have only one color. Grayscale images do not contain any information about color. Each pixel determines available different grey levels

A normal grayscale image contains 8 bits/pixel data, which has 256 different grey levels. In medical images and astronomy, 12 or 16 bits/pixel images are used.

3. Colour images

Colour images are three band monochrome images in which, each band contains a different color and the actual information is stored in the digital image. The color images contain gray level information in each spectral band

.The images are represented as red, green and blue (RGB images). And each color image has 24 bits/pixel means 8 bits for each of the three color band(RGB)

Fourier Transform (Computer Vision)

When we work in image processing, Fourier transform is an important image processing tool which is used to decompose an image into the frequency domain. the input image of the Fourier transform is the spatial domain(x,y) equivalent. the output of the Fourier transform represents the image in the frequency domain.

In the frequency domain image, each point represents a particular frequency contained in the spatial domain image. if an image has more high-frequency components (edges, stripes, corners), there will be a number of points in the frequency domain at high-frequency values.

.

Image Transformation

In image processing, image transformation can be defined as having control on its dimensional points to edit the images by moving them into three dimensional or two-dimensional space. Next in the article will perform some basic transformations of 2D images. Image Transformation works based on euclidean transformation, where the image can look shifted or rotated, but the structure of pixels and matrix will remain the same.

Roughly we can divide image transformation into two sections:

- Affine transformation.
- Projective transformation.

Affine transformation

As the name suggests in this transformation, preserving parallel relationships is one of the main concepts of this kind of transformation where lines will remain the same. Still, the square can change into a rectangle or parallelogram. It works by preserving the lengths and angles.

Projective Transformation:

As seen in the Affine transformation, we have less control in shifting the points, but in projective transformation, we have the freedom to move and shift the control points. It works on the projective view option where we see an object from its every plane. For example, a square image on paper from the front side looks like a square, but it will look like a trapezoid from the slight right or left side.

CONVOLUTION

Convolution is basically a filter that could be applied to images from the Computer Vision's perspective. Convolution in mathematical terms, is defined as an operator, which takes in two functions, and then produces a third function which best represents the amount of overlap between those two functions, but in a very different manner. Now you may be wondering how it works, and how it could apply a filter to an image. Convolution works by determining the value of the central pixel, by summing up the weighted values of all of the neighbor points together, and then creates a modified filtered image.

Another way of dealing images

Here we are going to discuss another method of dealing with images. This other method is known as convolution. Usually the black box(system) used for image processing is an LTI system or linear time invariant system. By linear we mean that such a system where output is always linear , neither log nor exponent or any other. And by time invariant we means that a system which remains same during time.

So now we are going to use this third method. It can be represented as.

It can be mathematically represented as two ways

g(x,y) = h(x,y) * f(x,y)

It can be explained as the "mask convolved with an image"

.

Or

g(x,y) = f(x,y) * h(x,y)

It can be explained as "image convolved with mask".

There are two ways to represent this because the convolution operator(*) is commutative. The h(x,y) is the mask or filter

.

What is mask?

Mask is also a signal. It can be represented by a two dimensional matrix. The mask is usually of the order of 1x1, 3x3, 5x5, 7x7 . A mask should always be in odd number, because other wise you cannot find the mid of the mask. Why do we need to find the mid of the mask. The answer lies below, in topic of, how to perform convolution?

How to perform convolution?

In order to perform convolution on an image, following steps should be taken.

- Flip the mask (horizontally and vertically) only once
- Slide the mask onto the image.
- Multiply the corresponding elements and then add them
- Repeat this procedure until all values of the image has been calculated.

Example of convolution

Let’s perform some convolution. Step 1 is to flip the mask.

Mask
Let’s take our mask to be this.

1	2	3
4	5	6
7	8	9

Flipping the mask horizontally

3	2	1
6	5	4
9	8	7

Flipping the mask vertically

9	8	7
6	5	4
3	2	1

Image
Let’s consider an image to be like this

2	4	6
8	10	12
14	16	18

Convolution

Convolving mask over image. It is done in this way. Place the center of the mask at each element of an image. Multiply the corresponding elements and then add them , and paste the result onto the element of the image on which you place the center of mask.

The box in red color is the mask, and the values in the orange are the values of the mask. The black color box and values belong to the image. Now for the first pixel of the image, the value will be calculated as

First pixel = (5*2) + (4*4) + (2*8) + (1*10)

= 10 + 16 + 16 + 10

= 52

Place 52 in the original image at the first index and repeat this procedure for each pixel of the image.

Why Convolution

Convolution can achieve something, that the previous two methods of manipulating images can't achieve. Those include the blurring, sharpening, edge detection, noise reduction e.t.c.