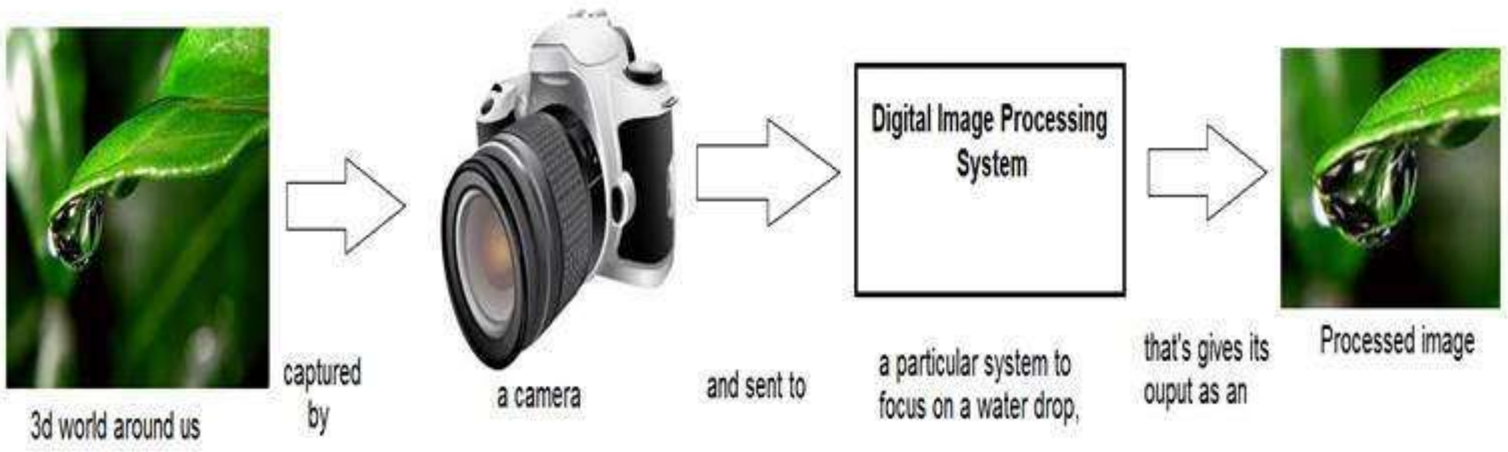


FUNDAMENTALS OF DIGITAL IMAGE PROCESSING

- Pre-requisites- Basic knowledge of programming, linear algebra, data structures and algorithms are required
- Nowadays image processing is becoming an important assisting tool in many branches of science such as computer science, electrical and electronic engineering, robotics, physics, chemistry, environmental science, biology, and psychology.

COURSE LEARNING OBJECTIVES (CLOs)

- Understand the basic concepts of digital image processing
- Understand fundamental image enhancement techniques on raw images
- Familiarize with various image restoration techniques

COURSE OUTCOMES(COs)

- Explain fundamental concepts of digital image processing
- Apply various techniques to enhance the image quality
- Analyse various image transformation techniques
- Implement various image processing techniques using MATLAB

SEMESTER – V			
Course Name	: Fundamentals of Digital Image Processing	Course Code:	20CI553
No. of Lecture Hours / Week	: 03	CIE Marks:	50
No. of Tutorial / Practical Hours / Week	: 00	SEE Marks:	50
Total No. of Lecture + Tutorial / Practical Hours	: 40	SEE Duration:	03 hr.
L: T: P	: 3:0:0	CREDITS:	03

Textbooks

1. Fundamentals of Digital Image Processing, A Practical Approach with Examples in MATLAB, Chris Solomon, U K Toby, Wiley-Blackwell, 2018

Reference Books

1. Rafael C. Gonzales, Richard E. Woods, Digital Image Processing, Third Edition, Pearson Education, 2010
2. William K Pratt, Digital Image Processing, John Willey, 2002
3. Anil Jain K, Fundamentals of Digital Image Processing, PHI Learning Pvt. Ltd., 2011

- Image processing is the process of transforming an image into a digital form and performing certain operations to get some useful information from it.
- It is also used to enhance the images, to get some important information from it. A certain number of algorithms are used in image processing.
- Image processing mainly include the following steps:

1.Importing the image via image acquisition tools;

2.Analysing and manipulating the image;

3.Output in which result can be altered image or a report which is based on analysing that image.

- For Example- MATLAB, Adobe Photoshop etc

Fundamental Image Processing Steps

Image Acquisition

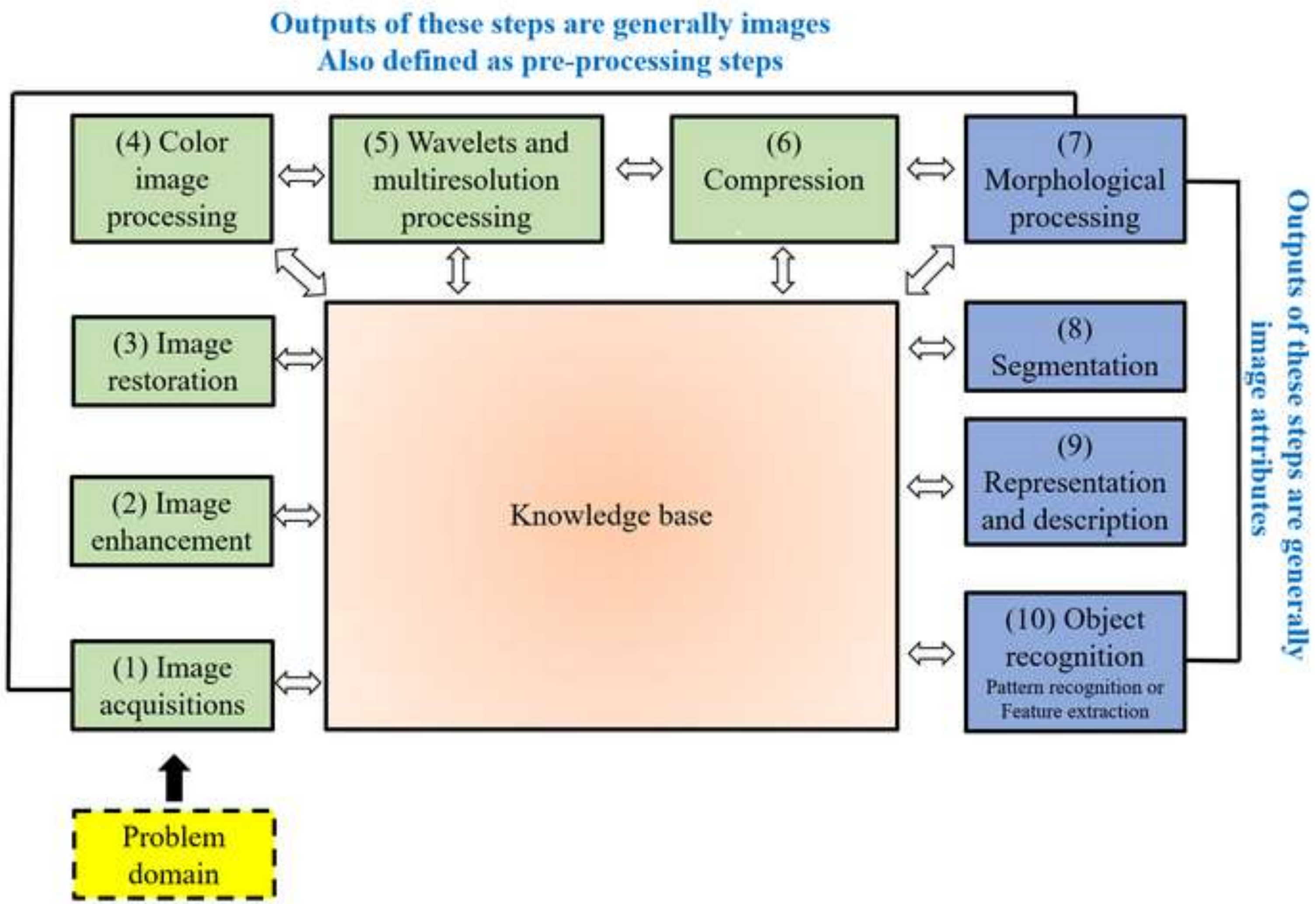
- Image acquisition is the first step in image processing. This step is also known as pre-processing in image processing. It involves retrieving the image from a source, usually a hardware-based source.

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.



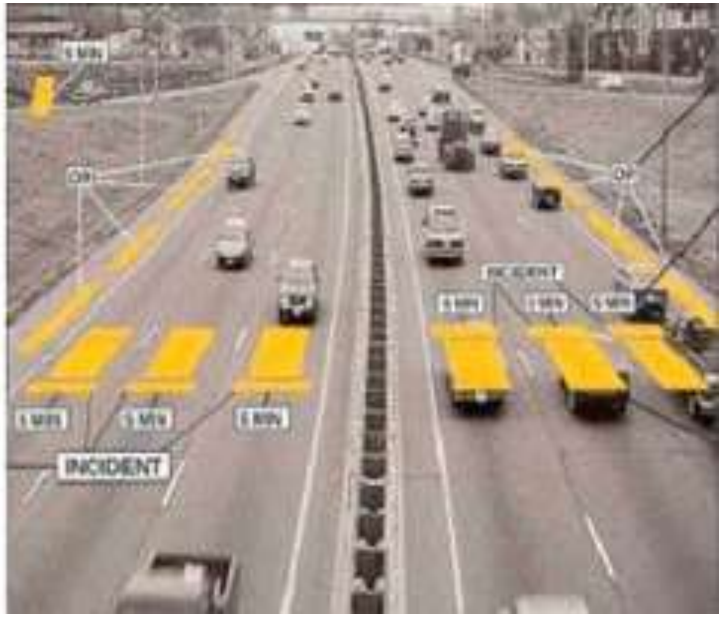
Outputs of these steps are generally images
Also defined as pre-processing steps
(4) Color image processing
(5) Wavelets and multiresolution processing
(6) Compression
(7) Morphological processing
(3) Image restoration
(8) Segmentation
(2) Image enhancement
Knowledge base
(9) Representation and description
(1) Image acquisitions
(10) Object recognition
Pattern recognition or Feature extraction
Problem domain
Outputs of these steps are generally image attributes

Applications of Image Processing

- Medical Image Retrieval- Image processing has been extensively used in medical research and has enabled more efficient and accurate treatment plans. For example, it can be used for the early detection of lung cancer using a sophisticated nodule detection algorithm in lung scans.

- **Traffic Sensing Technologies**

Normal traffic image



A VIPS image with detection zones

Image Reconstruction

- Reconstructing damaged images using image processing.

Face Detection

- Face detection is a vital tool used in security, biometrics and even filters available

  On most social media apps these days.

Module 1: Introduction

What is an Image?

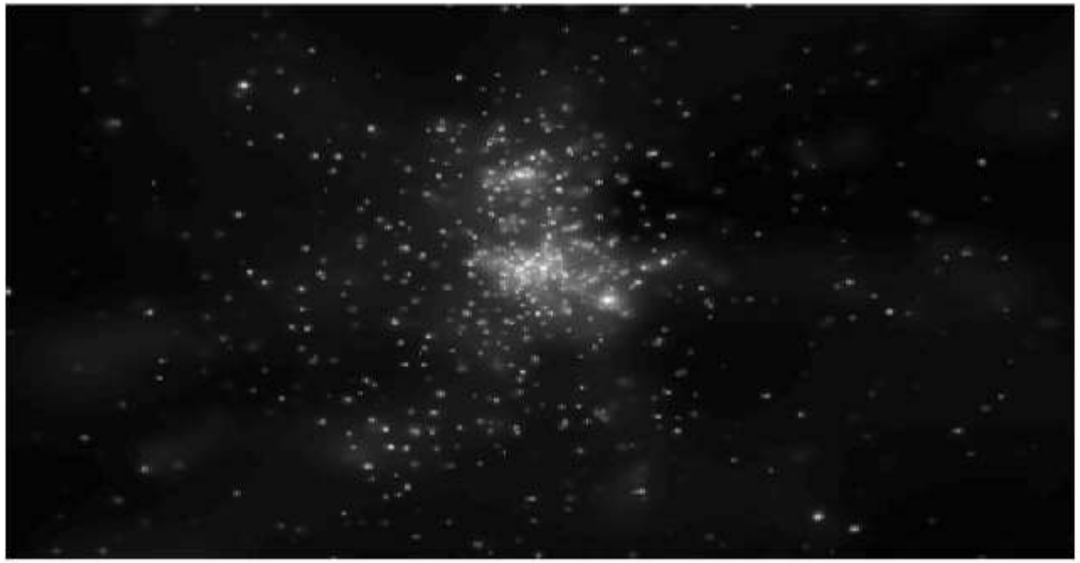
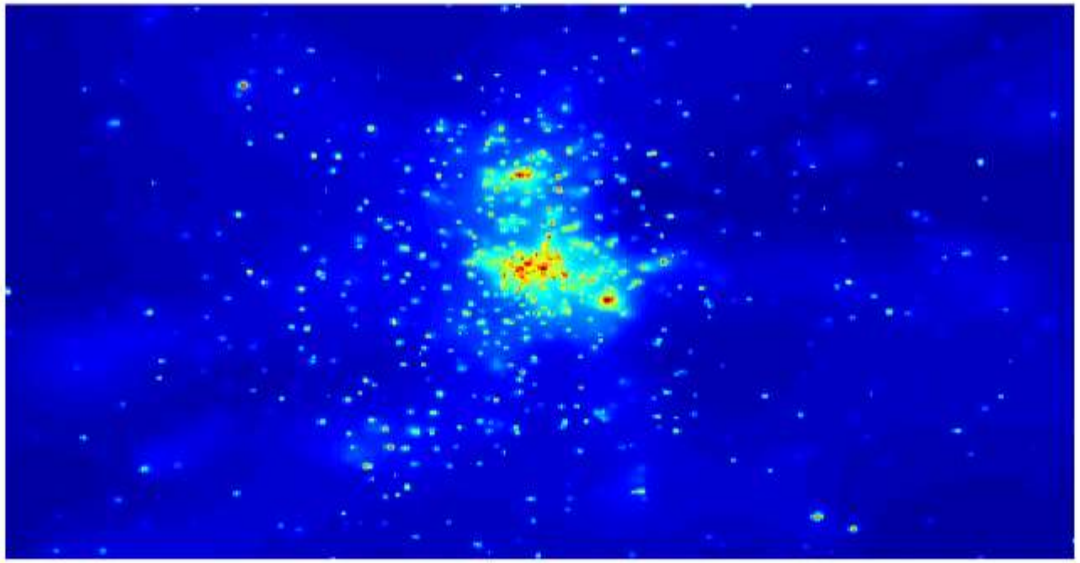
- A digital image is a discrete representation of data possessing both spatial (layout) and intensity (colour) information. An image is a multidimensional signal.

- The indices m and n respectively designate the rows and columns of the image. The individual picture elements or pixels of the image are thus referred to by their 2-D (m, n) index.

- This allows us to make more natural use of the powerful techniques of integral and differential calculus to understand properties of images and to effectively manipulate and process them.

Image colour

- An image contains one or more colour channels that define the intensity or colour at a particular pixel location I(m, n).

Example of grayscale (left) and false colour (right) image display

- A colour map assigns a specific shade of colour to each numerical level in the image to give a visual representation of the data.(all shades of grey from black (zero) to white (maximum)).

- In this example the jet colour map (as defined in Matlab) has been used to highlight the structure and finer detail of the image to the human viewer using a linear colour scale ranging from dark blue (low intensity values) to dark red (high intensity values).

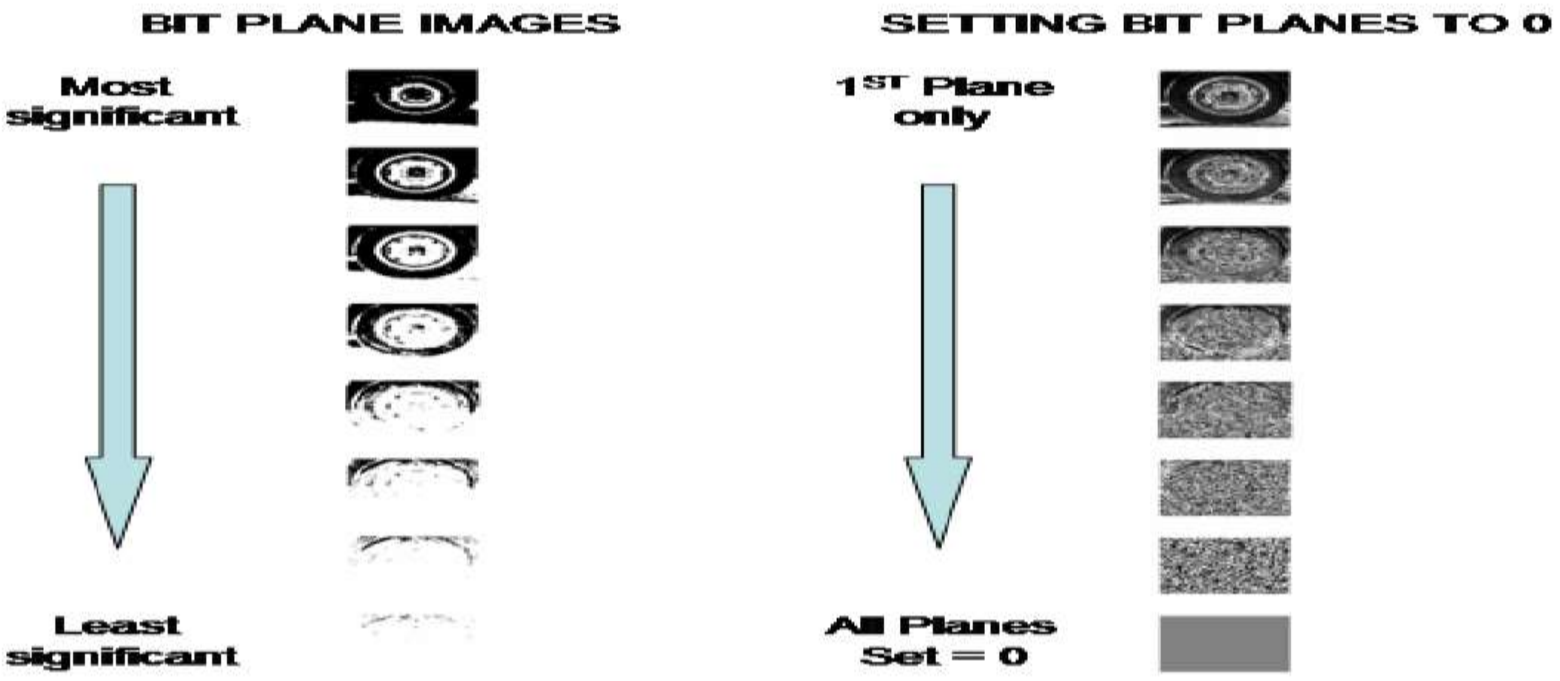
- In addition to greyscale images where we have a single numerical value at each pixel location, we also have true colour images where the full spectrum of colours can be represented as a triplet vector typically the (R,G,B) components at each pixel location.

- Other representations of colour are also possible and used quite widely, such as the (H,S,V) (hue, saturation and value (or intensity)). In this representation, the intensity V of the colour is decoupled from the chromatic information, which is contained within the H and S components.

Resolution and quantization

- Each individual image pixel determines the spatial resolution and colour quantization of the image.
- The representational power (or size) of an image is defined by its resolution. The resolution of an image source (e.g. a camera) can be specified in terms of three quantities
- Spatial resolution: The column (C) by row (R) dimensions of an image define the number of pixels used to cover the visual space captured by the image. It is commonly quoted as C*R (e.g. 640*480, 800*600, 1024*768, etc.)
- Temporal resolution: For a continuous capture system such as video, this is the number of images captured in a given time period. It is commonly quoted in frames per second (fps), where each individual image is referred to as a video frame.
- Bit resolution: This defines the number of possible intensity/colour values that a pixel may have. For instance a binary image has just two colours (black or white), a grey-scale image commonly has 256 different grey levels ranging from black to white whilst for a colour image it depends on the colour range in use.

Bit-plane splicing



We show the bit planes of an 8-bit grey-scale image of a car tyre descending from the most significant bit to the least significant bit. We see that the two or three least significant bits do not encode much useful visual information (mostly noise).
RHS shows the effect on the original image of successively setting the bit planes to zero. In a similar fashion, we see that these last bits do not appear to encode any visible structure.

- So, only the five most significant bits will produce an image which is practically visually identical to the original. Such analysis could lead us to a more efficient method of encoding the image using fewer bits – a method of image compression.
- For an 8-bit image, a pixel value of 0 is represented as 00000000 in binary form and 255 is encoded as 11111111.
- In bit-plane slicing, we divide the image into bit-planes. This is done by first converting the pixel values in the binary form and then dividing it into bit planes.

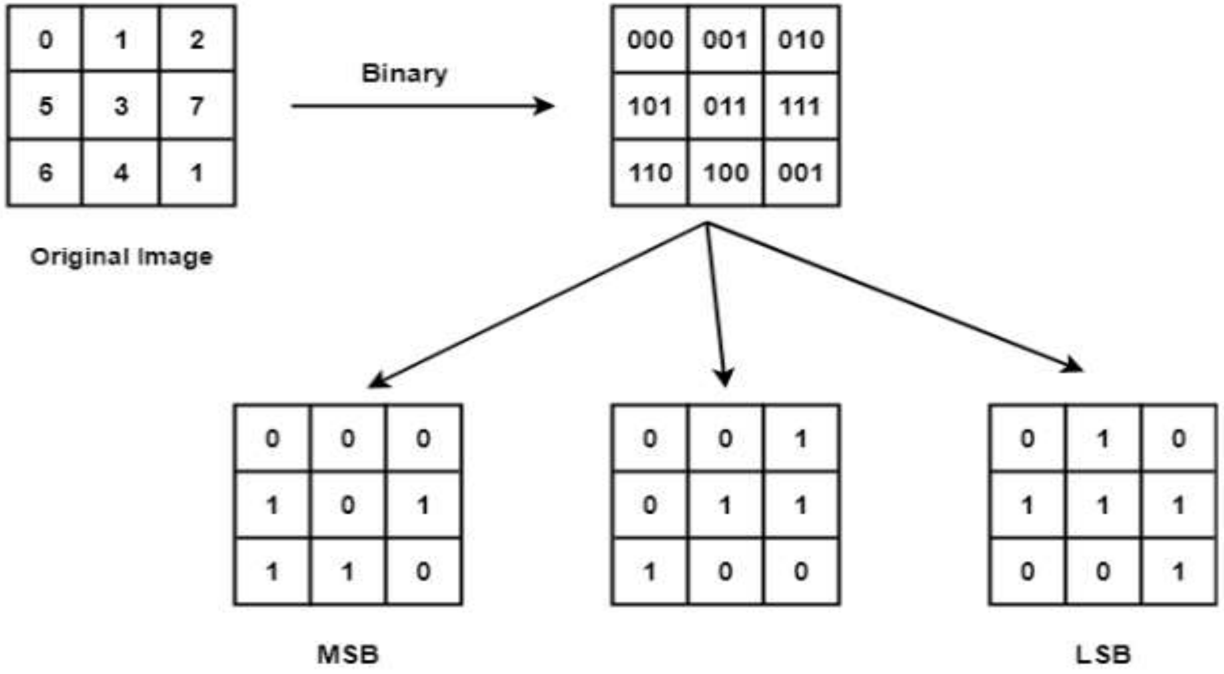


Image formats

Acronym	Name	Properties
GIF	Graphics interchange format	Limited to only 256 colours (8 bit); lossless compression
JPEG	Joint Photographic Experts Group	In most common use today; lossy compression; lossless variants exist
BMP	Bit map picture	Basic image format; limited (generally) lossless compression; lossy variants exist
PNG	Portable network graphics	New lossless compression format; designed to replace GIF
TIF/TIFF	Tagged image (file) format	Highly flexible, detailed and adaptable format; compressed/uncompressed variants exist

- GIF images are a very basic image storage format limited to only 256 grey levels.
- JPEG format is capable of storing up to a 24-bit RGB colour image, and up to 36 bits for medical/scientific imaging applications, and is most widely used for consumer-level imaging such as digital cameras.
- Bitmap format (BMP), originating in the development of the Microsoft Windows operating system.
- PNG format, designed as a more powerful replacement for GIF. TIFF, tagged image file format, represents an adaptable file format capable of storing a wide range of different image data forms.
- In general, photographic-type images are better suited towards JPEG or TIF storage, whilst images of limited colour/detail (e.g. logos, line drawings, text) are best suited to GIF or PNG (as per TIFF), as a lossless, full-colour format, is adaptable to the majority of image storage requirements.

Image data types

Binary images: are 2-D arrays that assign one numerical value from the set

{0, 1} to each pixel in the image. A fax (or facsimile) image is an example of a binary image.

Intensity or grey-scale images : are 2-D arrays that assign one numerical value to each pixel which is representative of the intensity at this point. As discussed previously, the pixel value range is bounded by the bit resolution of the image and such images are stored as N-bit integer images with a given format.

RGB or true-colour images : are 3-D arrays that assign three numerical values to each pixel, each value corresponding to the red, green and blue (RGB) image channel component, respectively. Conceptually, we may consider them as three distinct, 2-D planes so that they are of dimension C by R by 3, where R is the number of image rows and C the number of image columns. Commonly, such images are stored as sequential integers in successive channel order (e.g., $R_0G_0B_0$,$R_1G_1B_1$, ...) which are then accessed (as in Matlab) by I(C, R, channel) coordinates within the 3-D array.

Floating-point images: By definition, they do not store integer colour values. Instead, they store a floating-point number which, within a given range defined by the floating-point precision of the image bit resolution, represents the intensity. They may (commonly) represent a measurement value other than simple intensity or colour as part of a scientific or medical image.

Image: 24-bit RGB colour
Pixel Data Type: 3 x integer (0→255)
Image Format: JPEG

Image: 8-bit grayscale
Pixel Data Type: integer (0→255)
Image Format: GIF

Image: binary
Pixel Data Type: integer (0 or 1)
Image Format: PNG

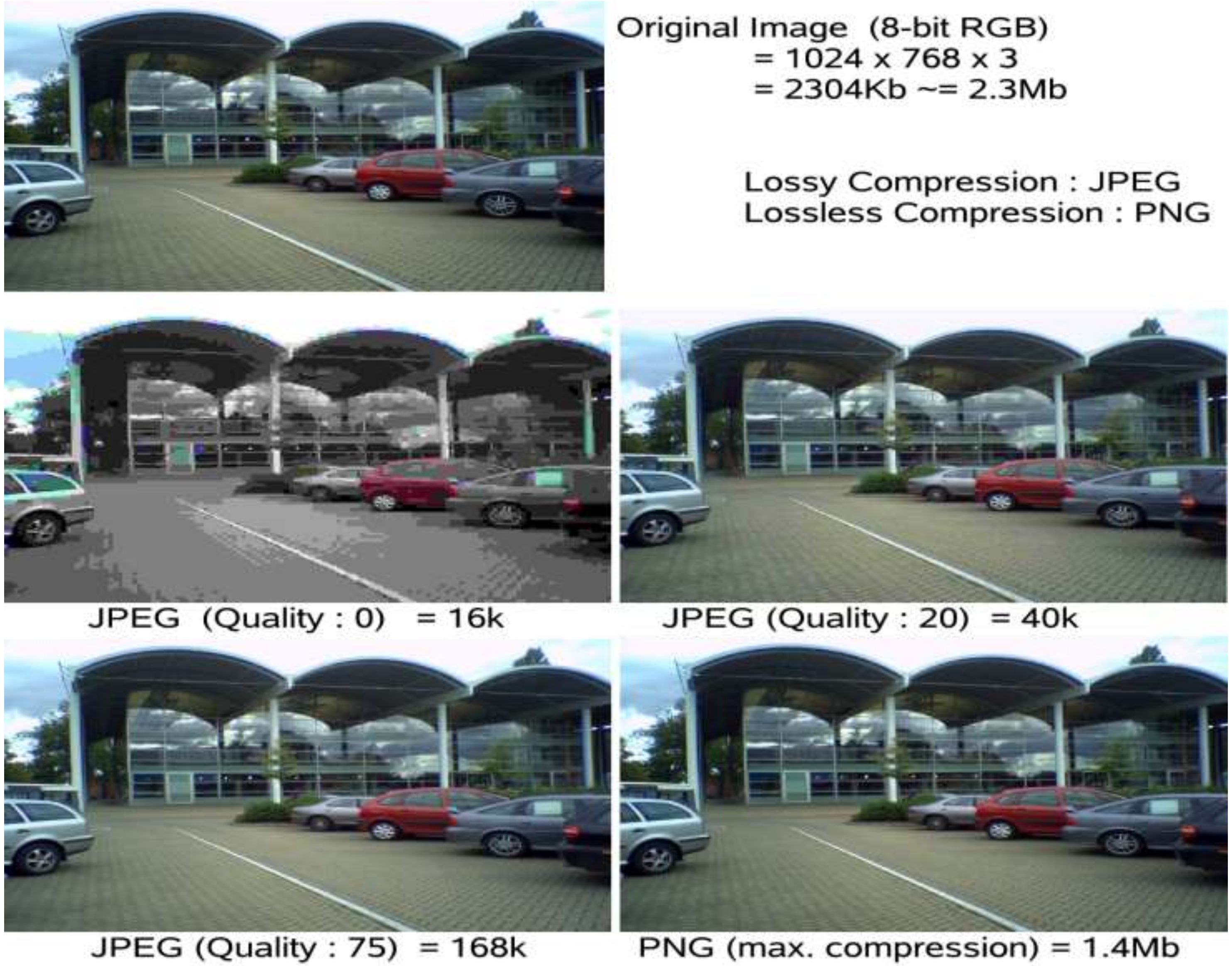
Image: floating point depth image
Pixel Data Type: floating point values
Image Format: TIFF

Image compression

- compressing an image can mean it takes up less disk storage and can be transferred over a network in less time.

Lossy compression

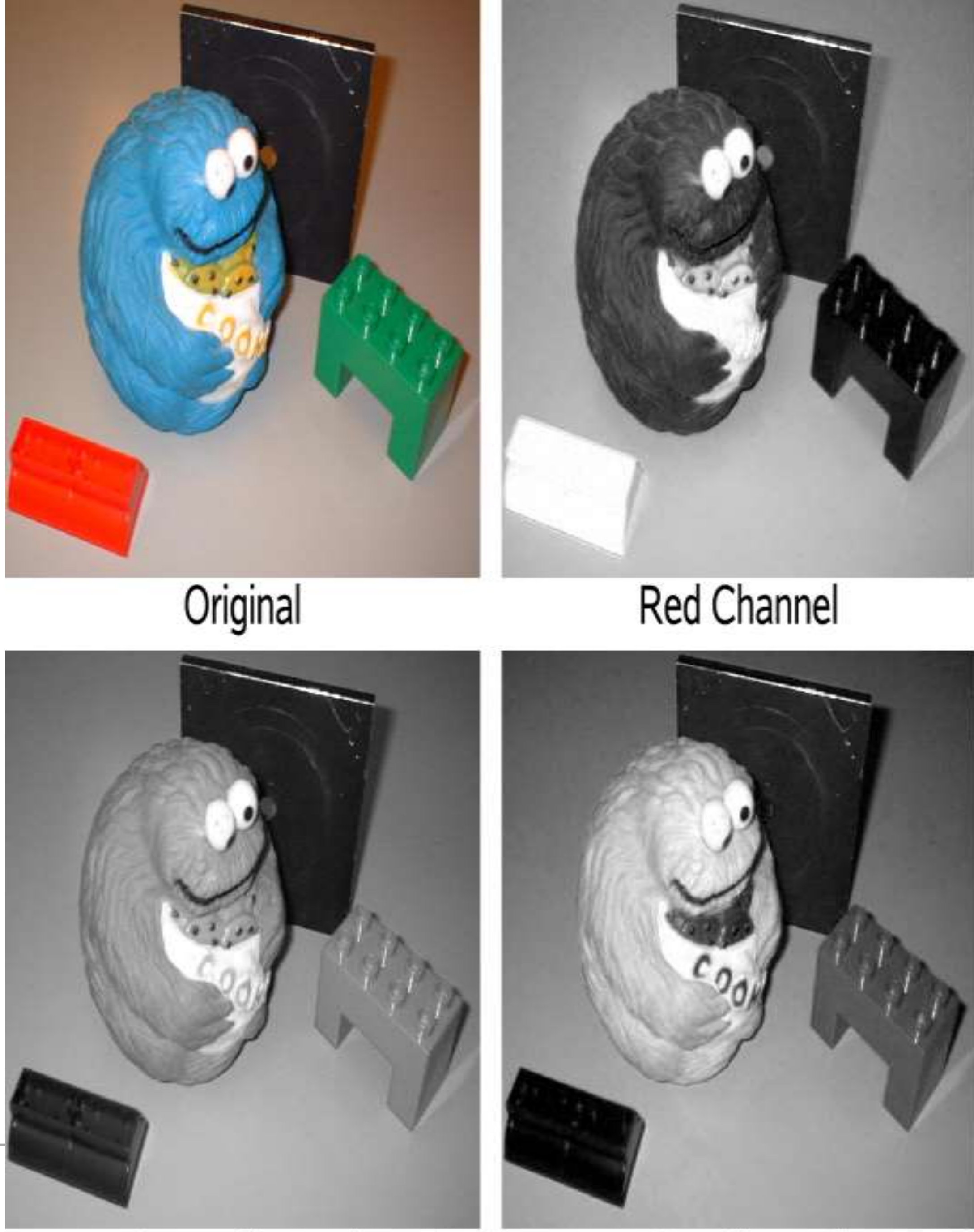
- operates by removing redundant information from the image. To remove some information from an image without any apparent change in its visual appearance.
- Storage of an image in one of the compressed formats employs various algorithmic procedures to reduce the raw image data to an equivalent image which appears identical (or at least nearly) but requires less storage.
- lossy compression techniques which reduce the storage volume at the expense of some loss of detail in the original image.



Example image compressed using lossless and varying levels of lossy compression

Colour spaces

- The representation of colours in an image is achieved using a combination of one or more colour channels that are combined to form the colour used in the image.
- The representation we use to store the colours, specifying the number and nature of the colour channels, is generally known as the colour space.
- An image is only a spatially organized set of numbers with each pixel location addressed as I(C,R)
- Binary images are 2-D arrays that assign one numerical value to each pixel.
- RGB or true-colour images are 3-D arrays that assign three numerical values to each pixel, each value corresponding to the red, green and blue component respectively.



Original

Red Channel

Green Channel

Blue Channel

RGB

- RGB colour space is essentially a 3-D colour space (cube) with axes R, G and B. Each axis has the same range 0->1 (this is scaled to 0–255 for the common 1 byte per colour channel, 24-bit image representation).
- The colour black occupies the origin of the

cube (position (0,0,0), corresponding to the

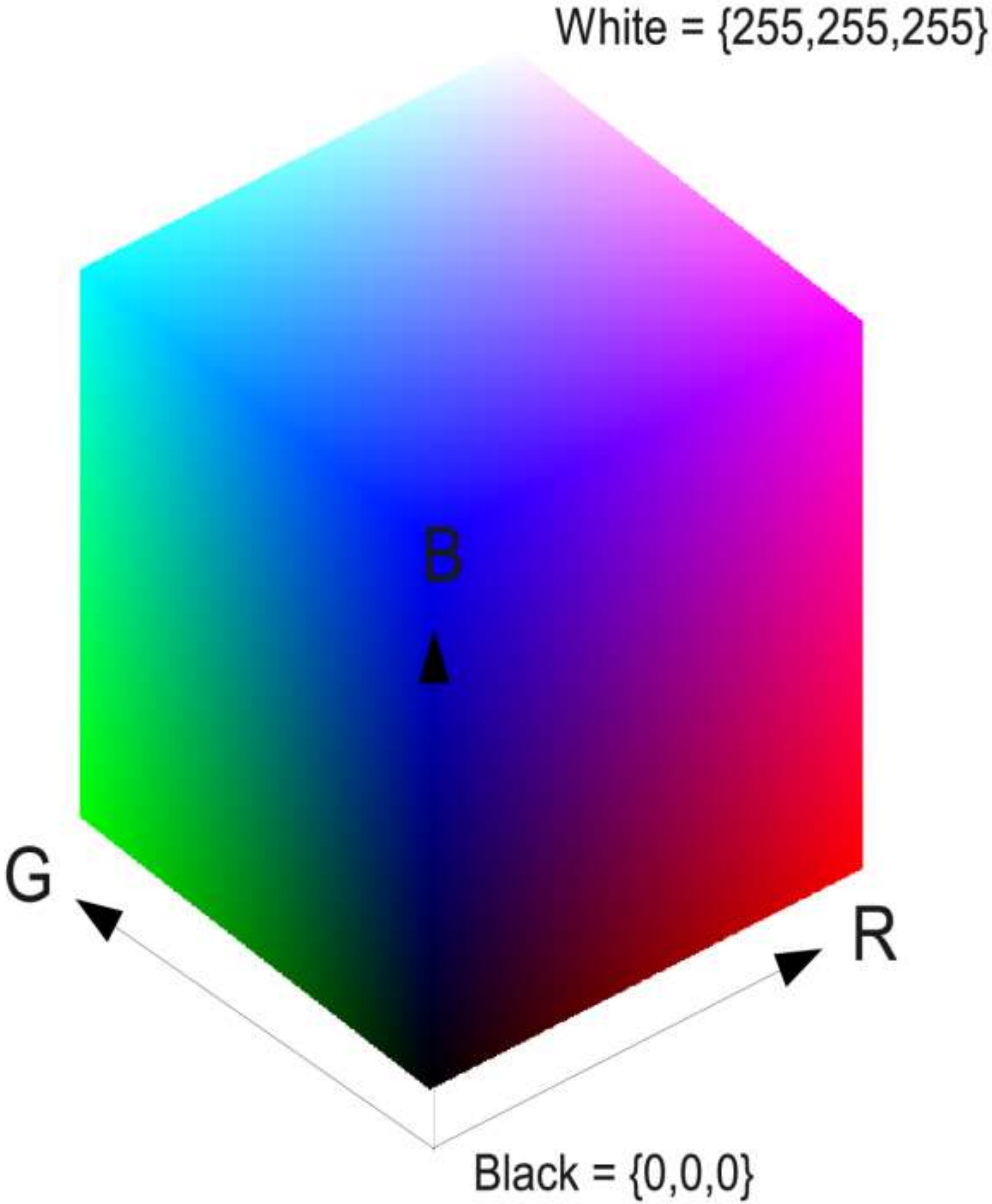
absence of all three colours; white occupies

the opposite corner (position(1,1,1)), indicating

the maximum amount of all three colours.

All other colours in the spectrum

lie within this cube.



RGB to grey-scale image conversion

- Grey-scale conversion is the initial step in many image analysis algorithms, as it essentially simplifies (i.e. reduces) the amount of information in the image.
- Although a grey-scale image contains less information than a colour image, the majority of important, feature related information is maintained, such as edges, regions, blobs, junctions.
- An RGB colour image, Icolour, is converted to grey scale, Igrey-scale, using the following transformation:

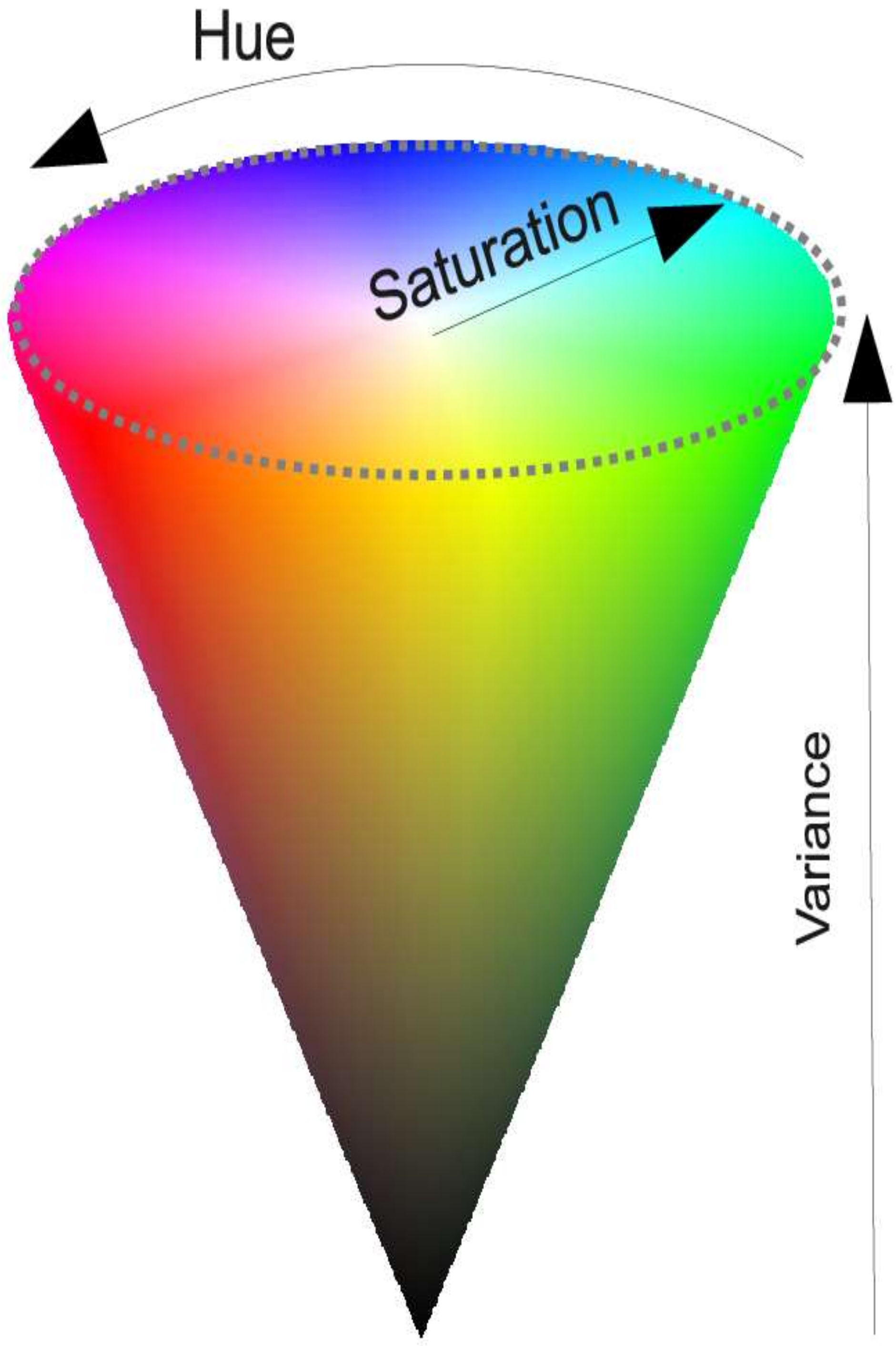
$$I_{\text{grey-scale}}(n,m) = \alpha I_{\text{colour}}(n,m,r) + \beta I_{\text{colour}}(n,m,g) + \gamma I_{\text{colour}}(n,m,b)$$

standardized weighting ensures uniformity (NTSC television standard)

$$\alpha = 0.2989, \beta = 0.5870 \text{ and } \gamma = 0.1140$$

Perceptual colour space

- An RGB image can be transformed into an HSV colour space representation as shown, Each of these three parameters can be interpreted as follows:
- . H (hue) is the dominant wavelength of the colour, e.g. red, blue, green
- . S (saturation) is the ‘purity’ of colour (in the sense of the amount of white light mixed
- with it)
- . V (value) is the brightness of the colour (also known as luminance).
- In the Matlab HSV implementation each of h, s and v are bounded within the range 0->1. For example, a blue hue may have a value of h=0.9, a saturation of s=0.5 and a value v=1 making it a vibrant, bright sky-blue.



- By examining the individual colour channels of images in the HSV space, we can see that image objects are more consistently contained in the resulting hue field than in the channels of the RGB representation, despite the presence of varying lighting conditions over the scene.

- As a result, HSV space is commonly used for colour-based image segmentation using a technique known as colour slicing.

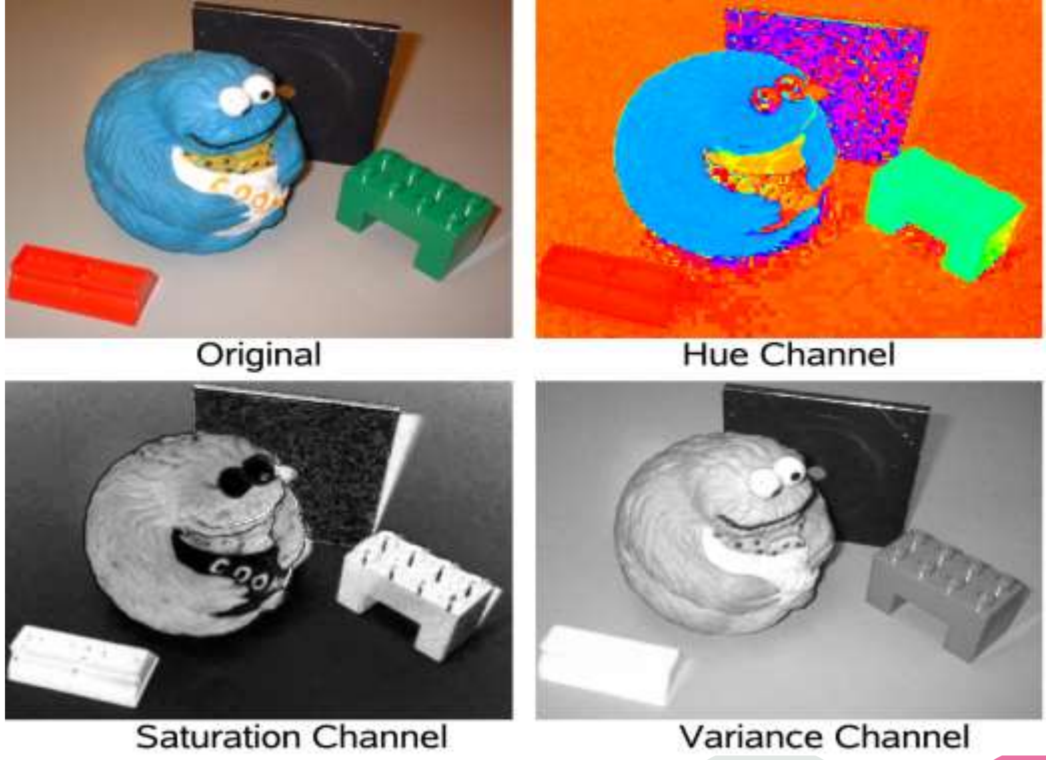


Image transformed and displayed in HSV colour space

Chapter 2: Formation

- How is an image formed? (Mathematical and engineering perspective)
- The image formation process can be summarized as a small number of key elements. In general, a digital image ‘s’ can be formalized as a mathematical model comprising a functional representation of the scene (the object function ‘o’) and that of the capture process (the point-spread function (PSF) p). Additionally, the image will contain additive noise n.

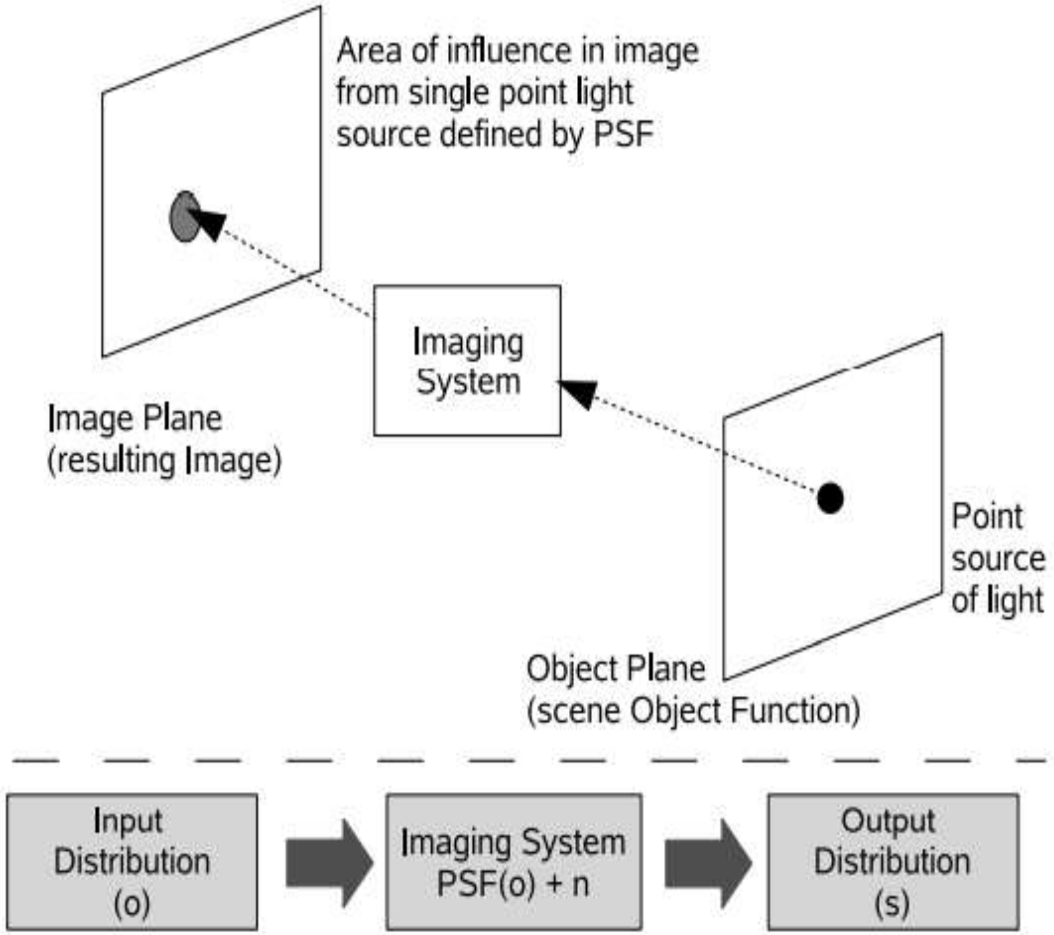
$$\text{Image} = \text{PSF} * \text{object function} + \text{noise}$$
$$s = p * o + n$$

- PSF-It is a characteristic of the imaging instrument (i.e., camera) and is a deterministic function (that operates in the presence of noise)
- Object function-This describes the object (or scene) that is being imaged and the way light is reflected from that structure to the imaging instrument.
- Noise-Noise is a stochastic function which is a consequence of all the unwanted external disturbances that occur during the recording of the image data.

- Convolution operator (*)-A mathematical operation which ‘smears’ (i.e. convolves) one function with another.

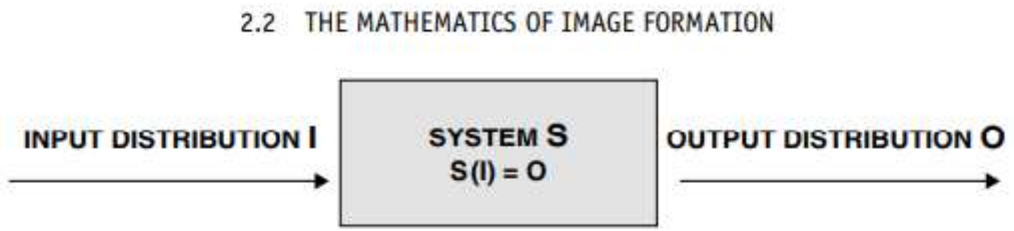
- Here, the function of the light reflected from the object/scene (object function) is transformed into the image data representation by convolution with the PSF.

- The PSF is a characteristic of the imaging instrument (i.e. camera). It represents the response of the system to a point source in the object plane, as shown in Figure, where we can also consider an imaging system as an input distribution (scene light) to output distribution (image pixels) mapping function consisting both PSF itself and additive noise as in Figure



The mathematics of image formation

- In a general mathematical sense, we may view image formation as a process which transforms an input distribution into an output distribution. Thus, a simple lens may be viewed as a ‘system’ that transforms a spatial distribution of light in one domain (the object plane) to a distribution in another (the image plane)



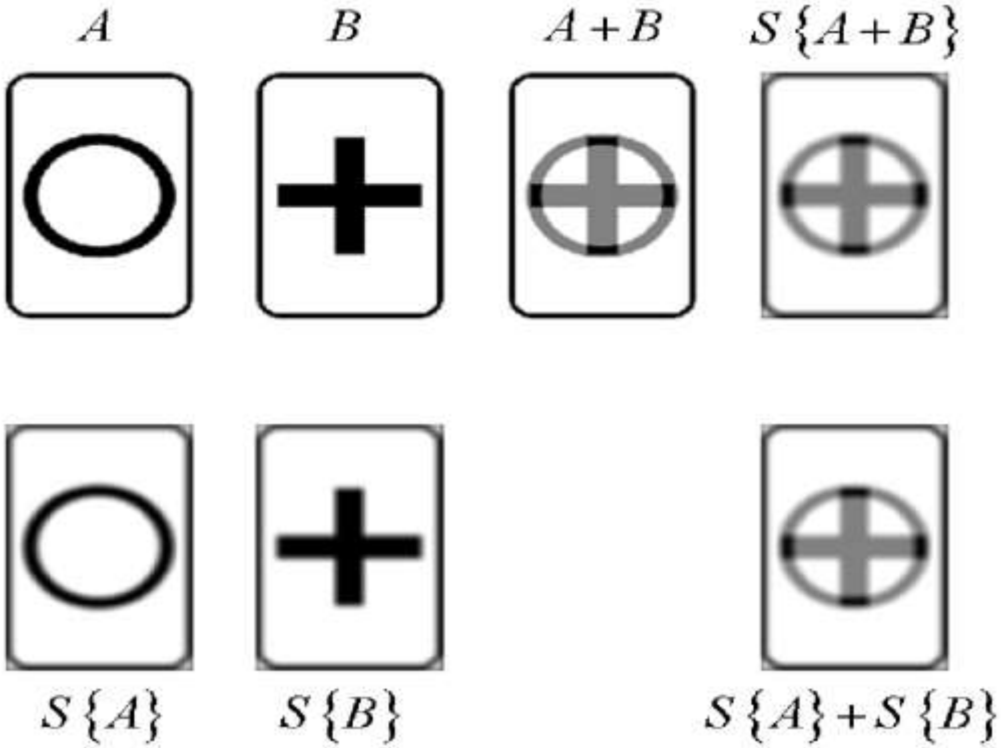
- Systems approach to imaging. The imaging process is viewed as an operator S which acts on the input distribution I to produce the output O
- Similarly, a medical ultrasound imaging system transforms a set of spatially distributed acoustic reflection values into a corresponding set of intensity signals which are visually displayed as a grey-scale intensity image.

Linear imaging systems

An imaging system described by operator S is linear if for any two input distributions X and Y and any two scalars a and b we have

$$S\{aX + bY\} = aS\{X\} + bS\{Y\}$$

- applying the linear operator to a weighted sum of two inputs yields the same result as first applying the operator to the inputs independently and then combining the weighted outputs.
- The first row of Figure shows the two input distributions, their sum and then the result of applying the linear operator (a blur) to the sum. The second row shows the result of first applying the operator to the individual distributions and then the result of summing them. In each case the final result is the same. The operator applied in Figure is a convolution with Gaussian blur.

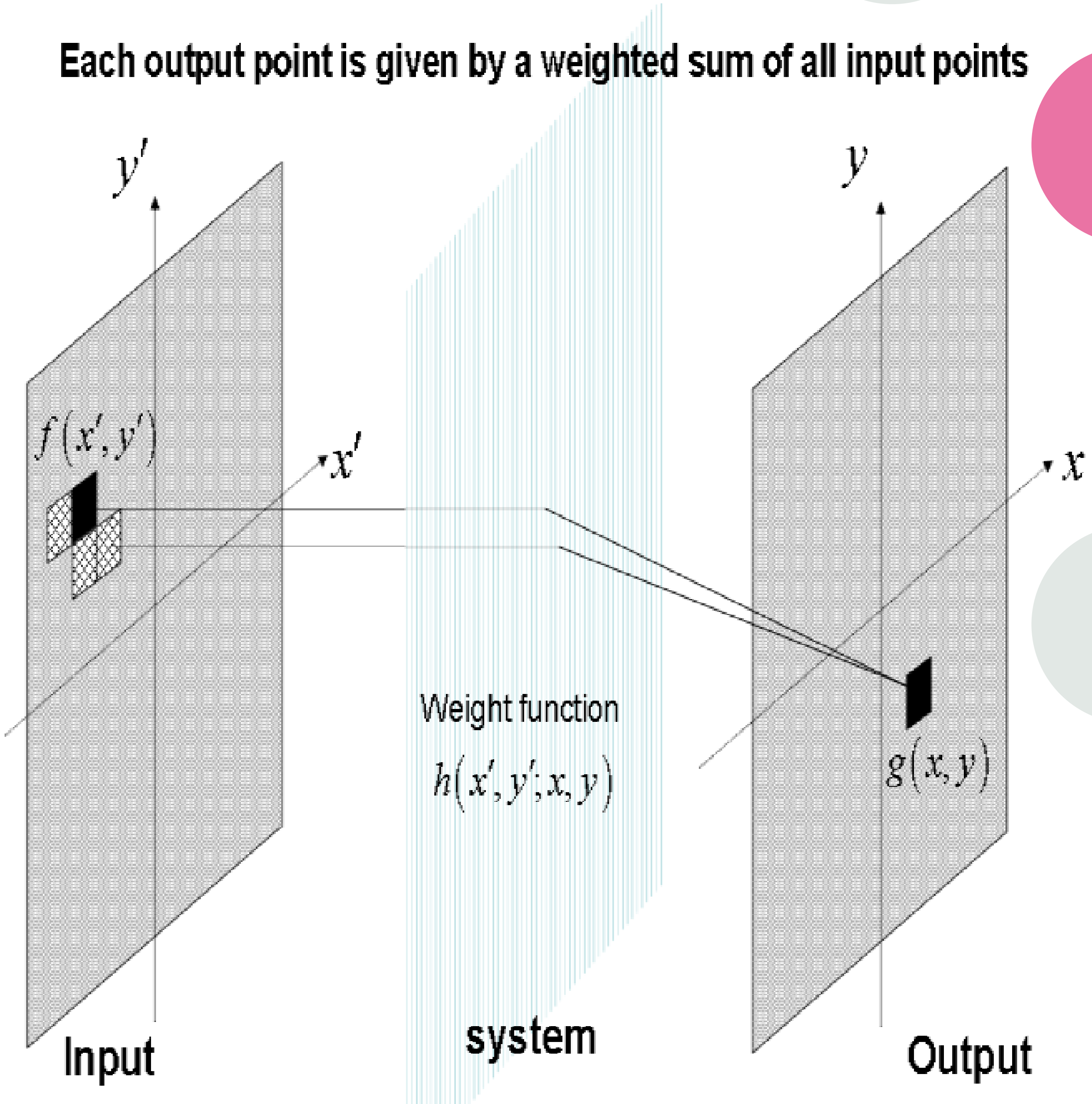


Linear superposition integral

Consider Figure, in which we have some general 2-D input function f(x’, y’) in an input domain (x’, y’) and the 2-D response g(x , y) of our imaging system to this input in the output domain (x , y). In the most general case, we should allow for the possibility that each and every point in the input domain may contribute in some way to the output. If the system is linear, however, the contributions to the final output must combine linearly. For this reason, basic linear image formation is described by an integral operator which is called the linear superposition integral.

$$g(x,y) = \iint f(x',y')h(x,y;x',y')\,\mathrm{d}x'\,\mathrm{d}y'$$

The output is given by a weighted sum (integral) of all input points. The weight function is called the point-spread function and specifies the contribution of each and every input point to each and every output point.



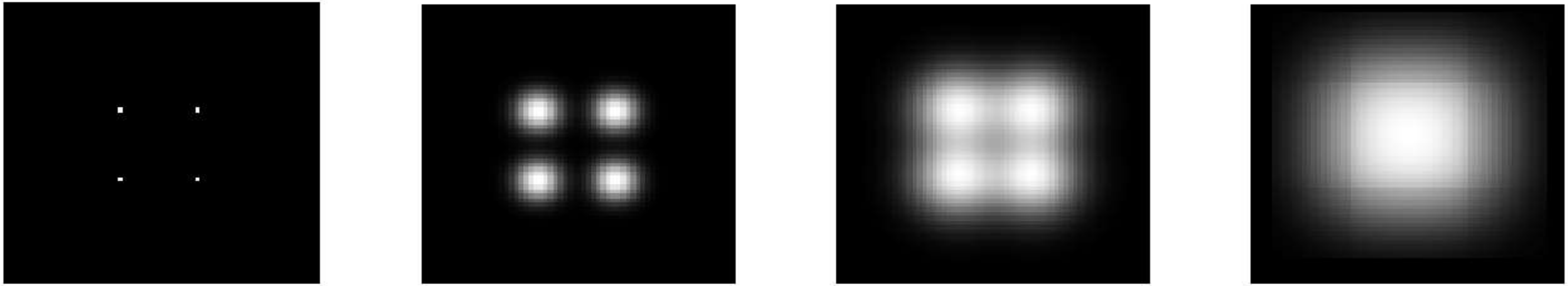
- . The linear superposition integral can be understood by breaking it down into three steps:

1) Take the value of the input function f at some point in the input domain (x’ ,y’) and multiply it by some weight h, with h determining the amount by which the input flux at this particular point contributes to the output point.
2) Repeat this for each and every valid point in the input domain multiplying by the appropriate weight each time.
3) Sum (i.e. integrate) all such contributions to give the response g(x, y)

- Clearly, it is the weighting function h which determines the basic behaviour of the imaging system. This function tells us the specific contribution made by each infinitesimal point in the input domain to each infinitesimal point in the resulting output domain.

The point-spread function

- Consider for a moment that any input distribution may be considered to consist of a very large (->∞) collection of very small (->infinitesimal) points of varying intensity.
- The PSF tells us what each of these points will look like in the output; so, through the linearity of the system, the output is given by the sum of the PSF responses.
- It is thus apparent that the PSF of such a linear imaging system (in the absence of noise) completely describes its imaging properties.
- A good or 'sharp' imaging system will generally possess a narrow PSF, whereas a poor imaging system will have a broad PSF



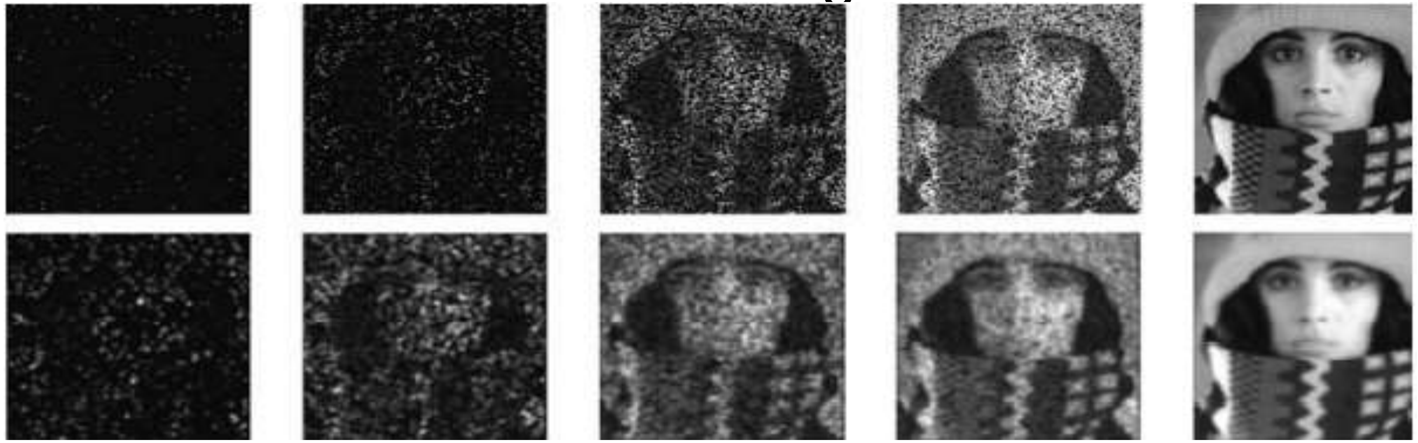
- The effect of the system PSF. As the PSF becomes increasingly broad, points in the original input distribution become broader and overlap

Linear shift-invariant systems and the convolution integral

- Convolution integrals are so important and common that they are often written in an abbreviated form:

$$g(x,y) = f(x,y) ** h(x,y) \quad \text{(2-D)}$$
$$g(x) = f(x) * h(x) \quad \text{(1-D)}$$

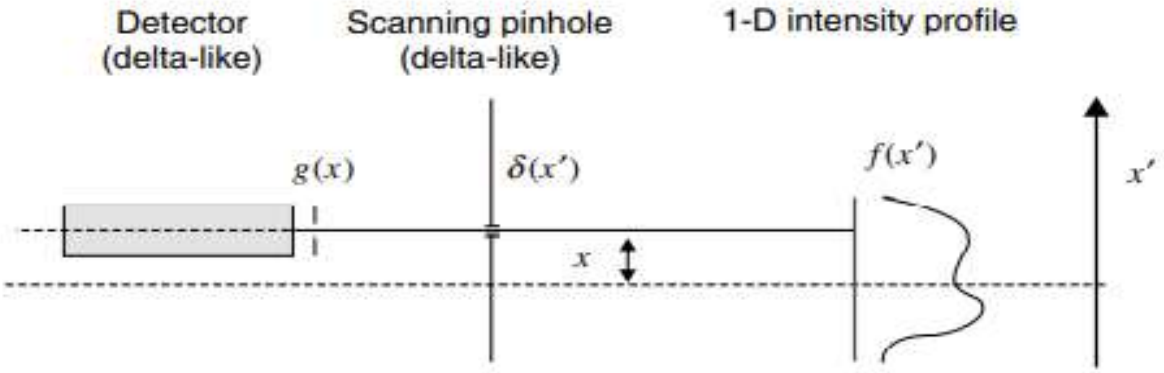
where the asterisks denote the operation of convolving the input function f with the system PSF h. In general, the function h in the convolution integrals above is called the kernel.



Each image in the top row shows arbitrary points in the input domain (increasing in number from left to right) whilst the images in the bottom row show the response of a LSI system to the corresponding input. The final image at the bottom right consists of a weighted superposition of responses to the input points, each of which has a similar mathematical shape or form.

Convolution: its importance and meaning

- A very large number of image formation processes are well described by the process of convolution. In fact, if a system is both linear and shift invariant, then image formation is necessarily described by convolution.
- The convolution theorem enables us to visualize and understand the convolution process in the spatial frequency domain.



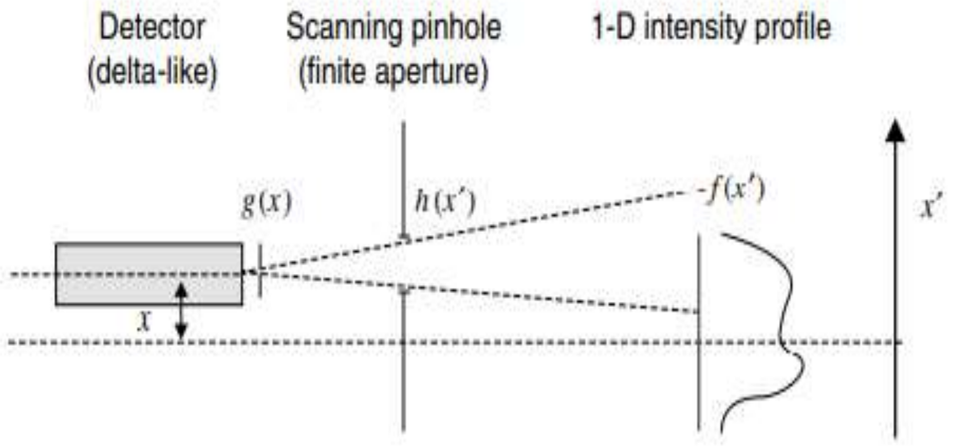
$$g(x)=\int f(x')\,h(x,x')\,dx'=\int f(x')\,\delta(x-x')\,dx'=f(x)$$

(a)

- A delta like detector views a source intensity profile through a delta-like scanning pinhole. The resulting image is a convolution of the source with the pinhole PSF. In principle, this replicates the source profile so that g(x) = f(x)

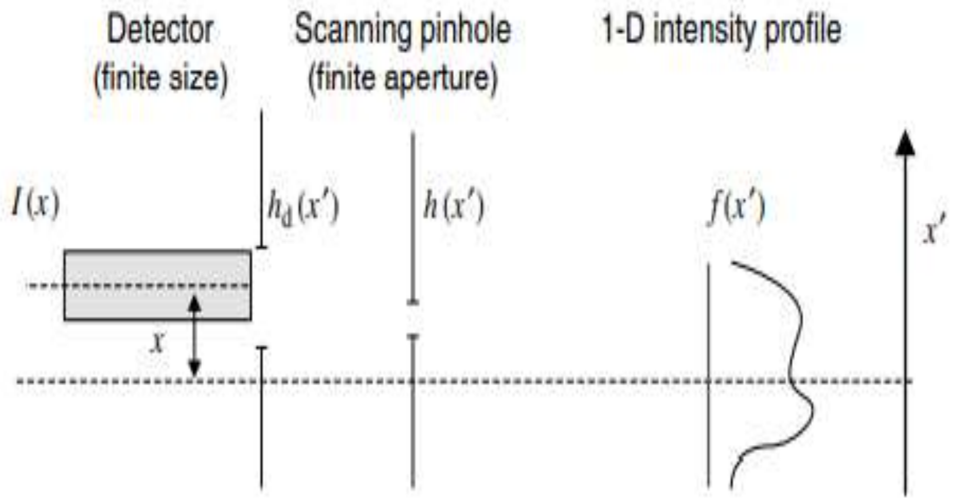
- Here, a delta like detector views a source intensity profile through a finite-size scanning pinhole. The resulting image is a convolution of the source with the pinhole PSF. All real systems exhibit such finite aperture effects.

- A finite-size detector viewing a source intensity profile through a finite scanning pinhole. The resulting image is a convolution of the source intensity with both the pinhole and detector PSFs. In general, if N LSI elements are involved in the image formation process, the resulting image is described by convolving the source with each corresponding PSF



$$g(x) = \int f(x')\, h(x, x')\, dx' = \int f(x')\, h(x - x')\, dx'$$

(b)



$$I(x) = \int g(x')\, h_d(x, x')\, dx' \text{ where } g(x) = \int f(x')\, \delta(x - x')\, dx'$$

$$I(x) = f(x') * h(x,) * h_d(x)$$

(c)

Multiple convolution: N imaging elements in a linear shift-invariant system

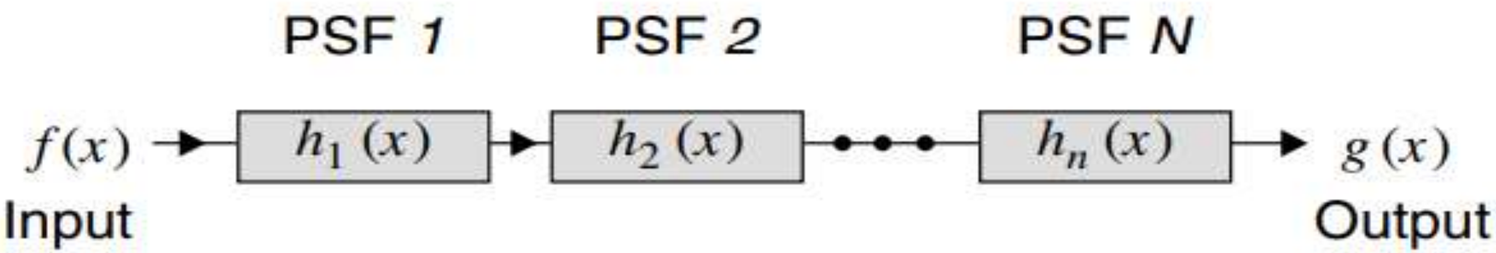
- we effectively have an imaging system in which two aperture functions describe the imaging properties: $h(x)$ and $h_d(x)$.
- we have seen that the recorded intensity $I(x)$ is given by successive convolutions of the input $f(x)$ with the PSF of the scanning aperture $h(x)$ and the PSF of the detector $h_d(x)$.

$$I(x) = f(x) * h(x) * h_d(x)$$

- Thus, in general, any processing sequence in which N linear and shift-invariant system elements act upon the input is described by a sequence of N convolutions of the input with the respective PSFs of the elements.

Digital convolution

- In digital image processing, signals are discrete not continuous. Under these conditions, convolution of two functions is achieved by discretizing the convolution integral.

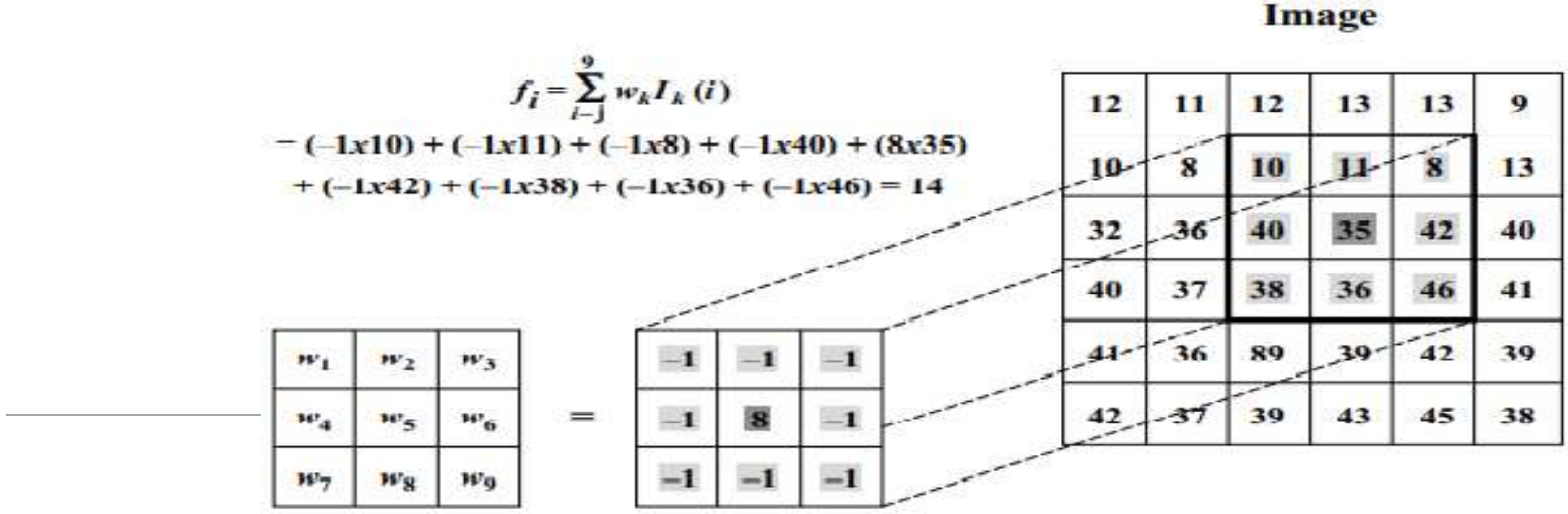


$$g(x) = f(x) * h_1(x) * h_2(x) \cdots * h_2(x)$$

- The output of an LSI system characterized by N components each having a PSF $h_i(x)$ is the repeated convolution of the input with each PSF

Discrete convolution

- The centre pixel of the kernel and the target pixel in the image are indicated by the dar grey shading. The kernel is ‘placed’ on the image so that the centre and target pixel match.
- The filtered value of the target pixel is then given by a linear combination of th neighbourhood pixels, the specific weights being determined by the kernel values. In thi specific case the target pixel of original value 35 has a filtered value of 14



The engineering of image formation

- The engineering aspect of image formation is where we will study the practical imaging scenarios.

The Camera

- A camera image of a conventional scene is essentially a projection of the 3-D world (i.e. the scene) to a 2-D representation (i.e. the image).
- If we assume an object/scene is illuminated by a light source (possibly multiple sources), then some of the light will be reflected towards the camera and captured as a digital image.
- The camera projection model transforms 3-D world coordinates (X,Y,Z) to 2-D image coordinates (x, y) on the image plane.
- The spatial quantization of the image plane projection into a discretized grid of pixels in turn transforms the 2-D image coordinate on the image plane to a pixel position (c, r). The majority of scene images we deal with will be captured using a perspective camera projection.

- **Perspective projection:** It can be stated as follows,

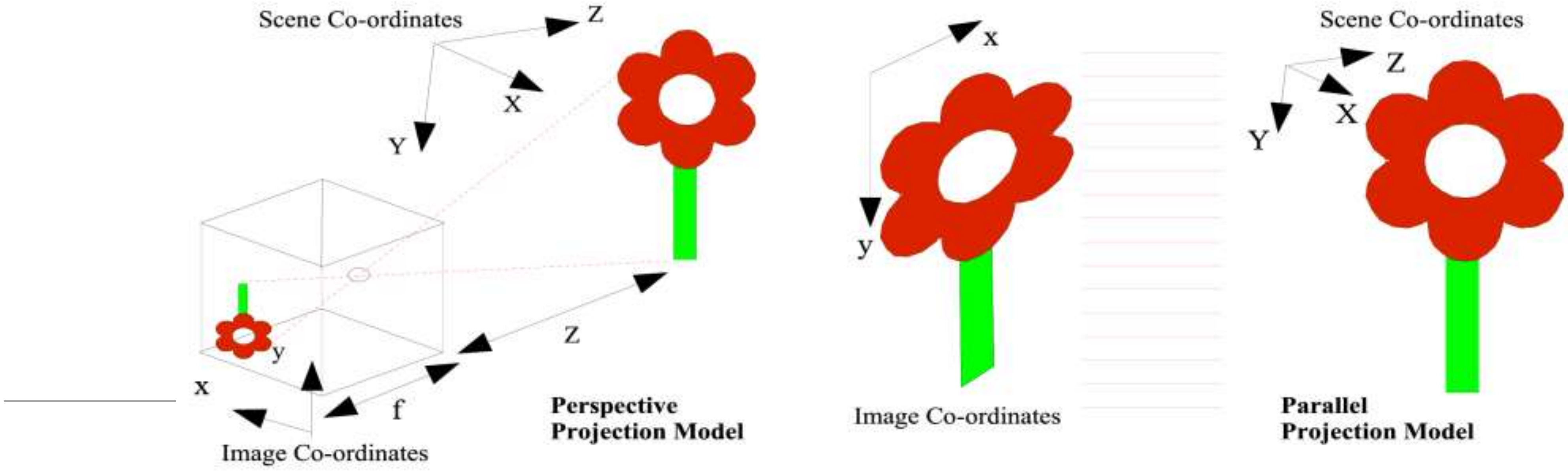
$$x = f\frac{X}{Z} \qquad y = f\frac{Y}{Z}$$

A position (X, Y, Z) in the scene (depth Z) is imaged at position (x, y) on the image plane determined by the focal length of the camera ‘f’ (lens to image plane distance). The Perspective projection has the following properties:

Foreshortening: The size of objects imaged is dependent on their distance from the viewer. Thus, objects farther away from the viewer appear smaller in the image.

Convergence: Lines that are parallel in the scene appear to converge in the resulting image. This is known as perspective distortion, with the point(s) at which parallel lines meet termed the vanishing point(s) of the scene.

- To be more precise this is the pin-hole perspective projection model, as we assume all the light rays pass through a single point and, hence, the scene is inverted (horizontally and vertically) on the image plane (Figure left).
- With a lens, the scene is still imaged upside down (and back to front), but this is generally dealt with by the camera and should be ignored for all applied purposes.
- It should be noted that the perspective projection is not easily invertible from an image alone – i.e. given 2-D image position (x, y) and camera focal length f it is not possible to recover (X,Y,Z).



- **Orthographic projection**: The orthographic (or parallel) projection is used by some specialist imaging instruments; for instance, a flat-bed scanner produces an orthographic projection of a scanned document, a medical scanner produces an orthographic projection of the human body. The orthographic projection is simply denoted as, $x=X \quad y=Y$

The digitization process:

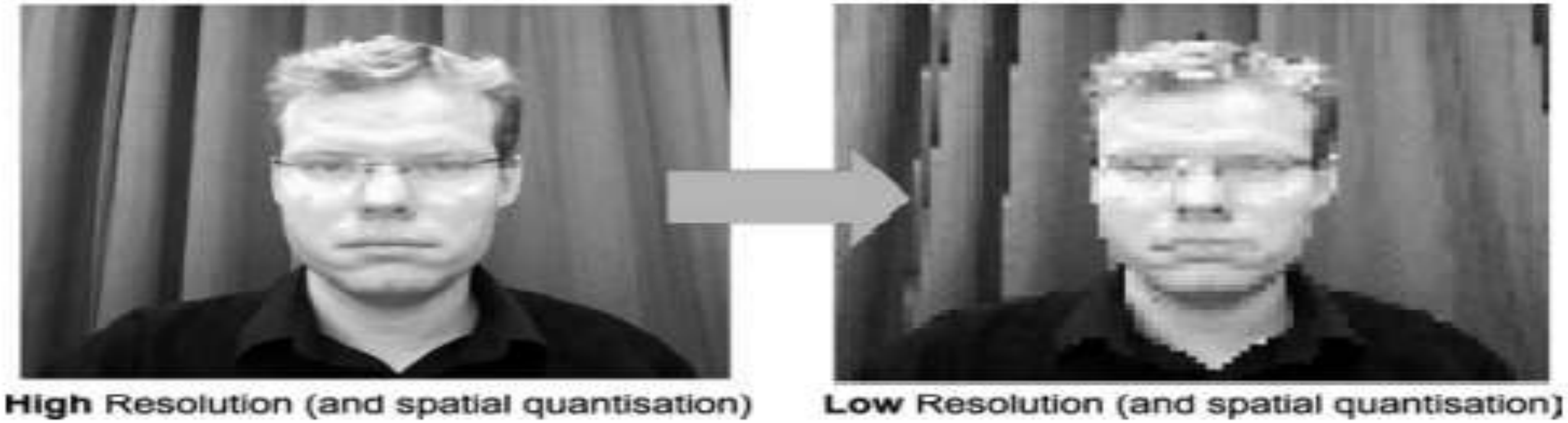
Discretization of the image into a finite-resolution image of individual pixels. The key concept in digitization is that of quantization: the mapping of the continuous signal from the scene to a discrete number of spatially organized points (pixels) each with a finite representational capacity (pixel depth)

Quantization:

Quantization in digital imaging happens in two ways: spatial quantization and colour quantization.

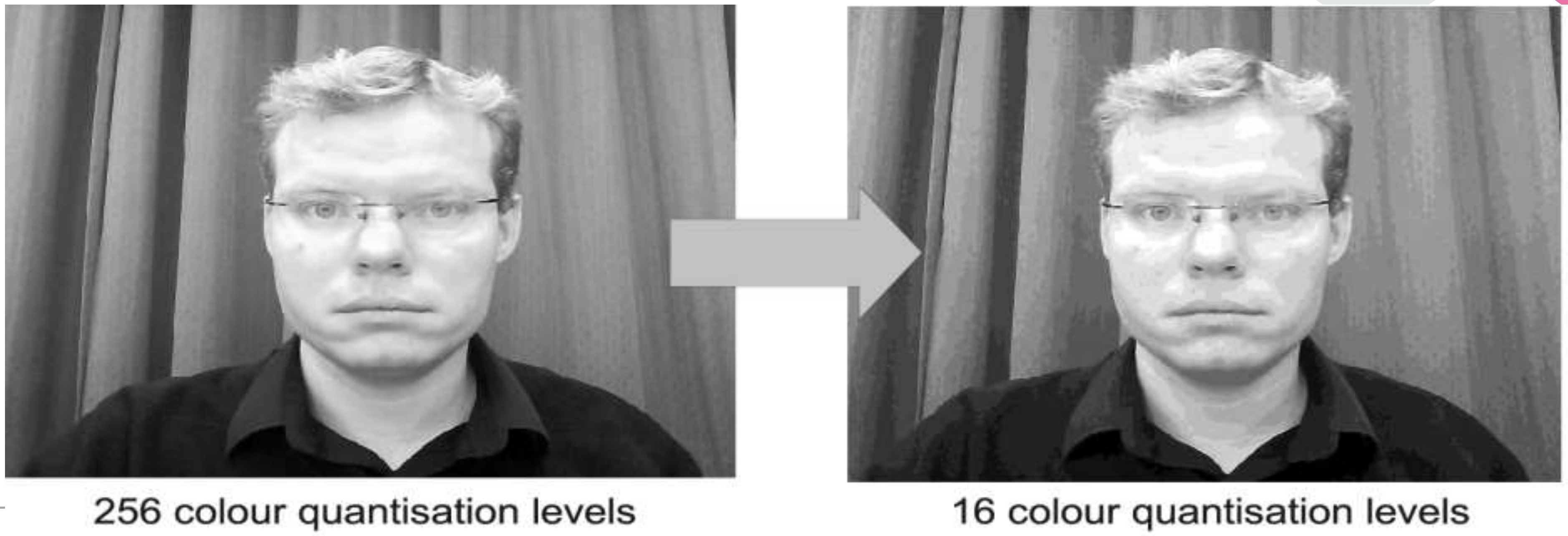
Spatial quantization corresponds to sampling the brightness of the image at a number of points. Usually a C*R rectangular grid is used but variations from rectangular do exist in specialist sensors. The quantization gives rise to a matrix of numbers which form an approximation to the analogue signal entering the camera. Each element of this matrix is referred to as a pixel – an individual picture element.

- Spatial quantization of the image causes aliasing to occur at edges and features within the image (Figure, left) and differing spatial quantization can affect the level of detail apparent within the image (Figure, right).



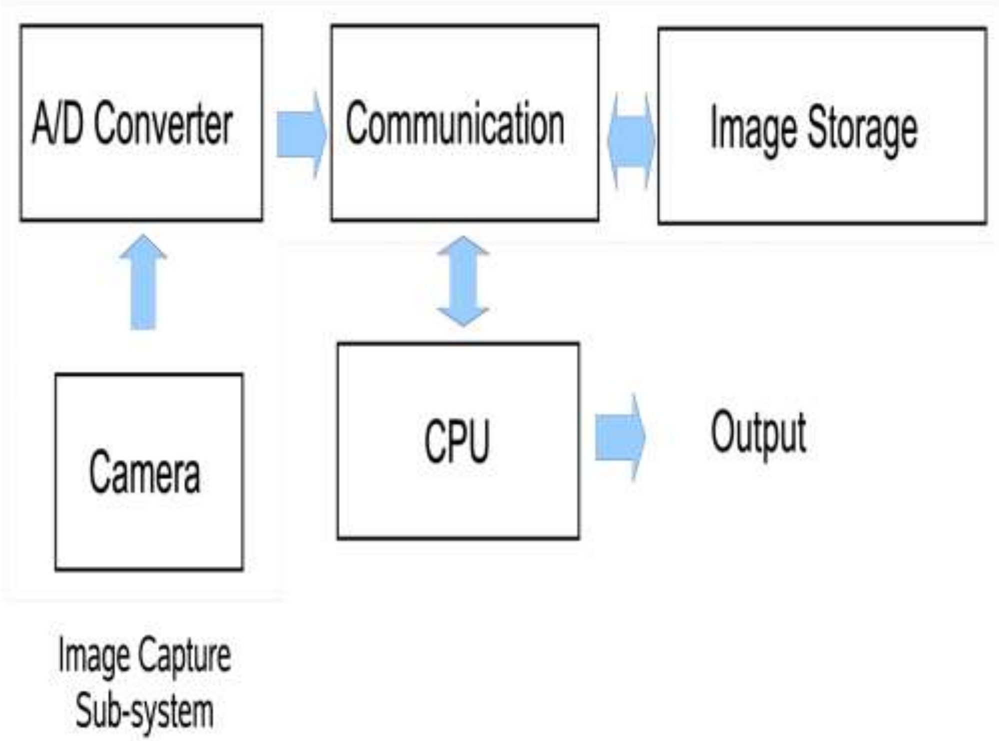
- **Intensity/color quantization**: Different levels of intensity quantization give different levels of image colour quality, as shown in Figure.
- For colour sensors each of the red, green and blue channels are similarly each quantized into an N-bit representation (typically 8-bits per channel, giving a 24-bit colour image). In total, 24-bit colour gives 16.7 million possible colour combinations. Although these representations may seem limited, given current computing abilities, it is worth noting that the human visual system can at most determine between ~40 different levels of grey and it is estimated to see around 7–10 million distinct colours.

- Note that.. although the human visual system cannot determine the difference between these levels, the same is not always true for image processing.
- A shallow colour depth (i.e. low-intensity quantization) can lead to different objects/features appearing at the same quantization level despite being visually distinct in terms of colour. Again, aliasing in intensity quantization can occur (e.g. the 16 grey-levels example in Figure 2.18)



Digitization hardware

- Traditionally, image capture was performed based on an analogue video-type camera connected to a computer system via an analogue to digital (A/D) converter.
- The converter required enough memory to store an image and a high-speed interface to the computer (commonly direct PCI or SCSI interface). When signalled by the computer, the frame grabber digitized a frame and stored the result in its internal memory.
- The data stored in the image memory (i.e. the digital image) can then be read into the computer over the communications interface. The schematic of such a system would typically be as per Figure 2.19, where we see an analogue link from the camera to the computer system itself and digitization is essentially internal to the hardware unit where the processing is performed.



Noise

- The main barrier to effective image processing and signal processing in general is noise. These factors cannot be (easily) controlled and thus introduce random elements into the processing pipeline. Noise is the key problem in 99 % of cases where image processing techniques either fail or further image processing is required to achieve the required result. As a result, a large part of the image processing domain, and any image processing system pipeline, is dedicated to noise reduction and removal.
- **Capture noise** can be the result of variations in lighting, sensor temperature, electrical sensor noise, sensor nonuniformity, dust in the environment, vibration, lens distortion, focus limitations, sensor saturation (too much light), underexposure (too little light)
- **Sampling noise** limitations in sampling and intensity quantization are a source of noise in the form of representational aliasing. The sampled digital image is not a true representation of the analogue image, but an alias of the original.

- **Processing noise** Limitations in numerical precision (floating-point numbers), potential integer overflow and mathematical approximations (e.g. pi= 3.142...) are all potential sources of noise in the processing itself.
- **Image-encoding noise** Many modern image compression techniques (e.g. JPEG used intrinsically by modern digital cameras) are lossy compression techniques. By lossy we mean that they compress the image by removing visual information that represents detail not general perceivable to the human viewer. The problem is that this loss of information due to compression undermines image-processing techniques that rely on this information. This loss of detail is often referred to by the appearance of compression artefacts in the image. In general,

  loss = compression artefacts = noise.
- **Scene occlusion** In the task of object recognition, objects are frequently obscured by other objects. This is known as occlusion. Occlusion poses a big problem for computer vision systems because you don't know what you cannot see. You may be able to infer recognition from a partial view of an object, but this is never as robust as full-view recognition. This limits available image information.

- **Salt and pepper noise :** This is caused by the random introduction of pure white or black (high/low) pixels into the image (Figure 2.21). This is less common in modern image sensors, although can most commonly be seen in the form of camera sensor faults (hot pixels that are always at maximum intensity or dead pixels which are always black). This type of noise is also known as impulse noise
- **Gaussian noise :** In this case, the random variation of the image signal around its expected value follows the Gaussian or normal distribution (Figure 2.21). This is the most commonly used noise model in image processing. This type of noise is also known as additive noise.

Original Image

Salt and Pepper (impulse) noise

Thank you