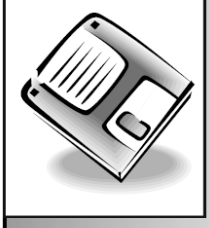
Summary



- Images are a way of recording and presenting information in a visual form.
- Digital image processing is a collection of techniques for the manipulation of digital images by computers.
- A digital image is composed of finite number of elements called pixels. Each pixel has a particular location and value.
- Sampling is the process of measuring the value of the physical image at discrete intervals in space. In order to sample the signal, it is necessary that the signal should be a band-limited signal.
- Each image sample corresponds to a small region of the physical image, which is called a picture element or pixel. A digital image corresponds to a two-dimensional array of pixels.
- Quantisation is the process of replacing the continuous values of the sampled image with a discrete set of quantisation levels. The number of quantisation levels employed governs the accuracy by which the image values are displayed.
- A vector image or geometric image represents an image mathematically through the use of geometrical primitives such as points, lines, curves and polygons.
- Image sensor is a 2D array of light-sensitive elements that convert photons to electrons. CCD and CMOS image sensors are widely used in image-capturing devices like digital cameras and camcorders.
- The resolution of a CCD image sensor is better than a CMOS image sensor, whereas a CMOS image sensor is cost-effective when compared to a CCD.
- Scanners convert analog images into digital image files. Common image scanners are drum scanners, flatbed scanners and flying-spot scanners.
- A file format is a method used to store digital data. The image file consists of two parts: (i) file header, and (ii) image data. The common image file formats include TIFF, BMP, JPEG, PNG, GIF and PSD.
- Popular image file formats such as TIFF, GIF, PNG and BMP use lossless compression algorithm whereas JPEG standard comes with both lossless and lossy compression algorithm.
- GIF comes in two formats: GIF87 and GIF89a. GIF89a supports simple animations.

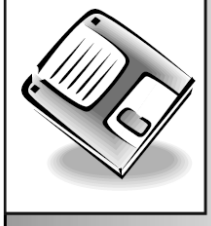
Summary

- Signals are variables that carry information. An example of a 2D signal is a still image.
- A 2D sequence is periodic if it repeats itself in a regularly spaced interval.
- A 2D sequence $x(n_1, n_2)$ is separable if it can be represented as $x(n_1, n_2) = x_1(n_1)x_2(n_2)$.
- A system can be viewed as an operation or a set of operations performed on the input signal to produce an output signal.
- A system is a shift-invariant system if its input–output characteristics do not change with time.
- A system is linear if it obeys the superposition theorem. The principle of superposition requires that the response of the system to a weighted sum of signals should be equal to the corresponding weighted sum of the output of the system to each of the individual input signals.
- Linear Shift Invariant (LSI) systems are uniquely represented by their 2D impulse response.
- A system is static or memoryless if its output at any instant depends at most on the input sample but not on the past and future samples of the input.
- A 2D LSI system is BIBO stable if its impulse response is absolutely summable.
- A Z-transform is widely used to analyse a 2D system. The Z-transform of a 2D sequence $x(n_1, n_2)$ is given by

$$X(z_1, z_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} x(n_1,n_2) z_1^{-n_1} z_2^{-n_2}$$

- Filters can be considered as systems that process the input signal and give an output signal. Filters can be broadly classified into (a) FIR filters, and (b) IIR filters. FIR filters are more popular since they exhibit linear phase characteristics.
- 2D FIR filters can be designed using windowing technique, frequency sampling technique and the optimal method.
- 2D IIR filters are recursive in nature. The stability of a 2D IIR filter is not always guaranteed.

Summary



- Convolution basically computes the new value of a pixel as a weighted sum of pixel values in a certain neighbourhood surrounding the pixel.
- Convolution can be used to carry out the linear filtering process.
- Linear and circular convolution can be computed through (i) graphical method, (ii) Z-transform method, and (iii) matrix method.
- The operation of convolution is not commutative.
- The computational load of convolution is reduced considerably if the utilised kernel is separable.
- Circular convolution can be used to perform interpolation of images.
- Correlation is commonly used to determine similarities between images or parts of images. It is possible to use correlation to determine the global displacement between two images of the same scene.
- Correlation can be broadly classified into (i) autocorrelation, and (ii) cross-correlation. If one finds the similarity of a signal to itself then it is autocorrelation. Cross-correlation is the similarity between two different signals.
- Autocorrelation is self-convolution without reflection of either function.
- The Fourier transform of the autocorrelation function is the power spectrum.
- Convolution in spatial domain is equal to multiplication in the frequency domain.

Summary

- A transform is basically a representation of a signal. A transform changes the representation of a signal by projecting it onto a set of basis functions. The transform does not change the information content present in the signal.
- Different types of image transforms are Fourier, Walsh, Hadamard, Slant, Cosine, Sine, KL, Radon and Singular Value Decomposition.
- The transformation matrix A is unitary if it obeys the following condition

$$A^{-1} = A^{*T}$$

- The transformation matrix A is orthogonal if it obeys the following relation

$$A^{-1} = A^{*T}$$

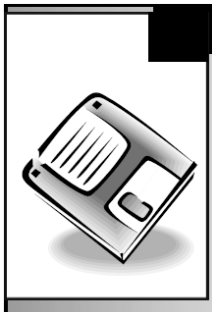
- The king of all transforms is Fourier transform. The 2D Fourier transform of a signal $f(m, n)$ is given by

$$F(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n) e^{-j\frac{2\pi}{M}nl} e^{-j\frac{2\pi}{N}mk}$$

- $e^{-j\frac{2\pi}{M}nl}, e^{-j\frac{2\pi}{N}mk}$ represents the Fourier basis. The significance of Fourier basis are the following
  1. Fourier basis is with both real and imaginary terms.
  2. Fourier basis will not lose its identity due to mathematical operations like addition, subtraction, differentiation and integration.
- A Fourier transform is represented by its magnitude and phase. The magnitude information gives how much of a certain frequency component is present and phase information gives where the frequency component is in the image.
- The number of sign changes along a column of the transform matrix is known as the sequency of that column.
- Fourier transform is separable. The separable property allows one to compute a two-dimensional DFT as two separate one-dimensional DFTs.
- Narrowing a function broadens its Fourier transform and vice versa, i.e., compression in one domain is equal to expansion in another domain.
- Spatial shift will result in phase shift (shifting property of Fourier transform)
- Rotating a function in the spatial domain will result in the rotation of the spectrum in the frequency domain.
- Convolution of two functions is similar to multiplication of their Fourier transforms.
- The Fourier transform of a Gaussian function is another Gaussian.
- A set of mutually orthonormal basis functions, with values +1 or –1 constitutes the Walsh transform kernels.
- The computation of Walsh coefficients is simple and it involves only addition and subtraction operations.
- The Hadamard matrix of order n satisfies the orthogonality condition $H_n \times H_n^T = nI_n$

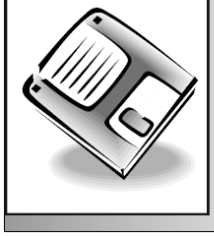
- A KL transform depends on the second-order statistics of the input data. It is an optimal transform with respect to energy compaction.
- Discrete Cosine Transform (DCT) is real and orthogonal. DCT has excellent energy compaction for highly correlated data. DCT is used in JPEG standard.
- Haar functions are non-sinusoidal orthogonal functions. The Haar transform is the simplest discrete wavelet transform.
- The slant transform is real and orthogonal. Slant transforms possess good energy compaction property.
- Some of the applications of image transform include filtering, restoration, compression, enhancement and image analysis.
- Many common unitary transforms tend to pack a large fraction of signal energy into a few transform coefficients which is generally termed energy compaction.

Summary



- The goal of image enhancement is to improve the perception of the image through modification of intensity functions, image spectral content, or a combination of these functions.
- Removing blurring and noise, increasing contrast, and revealing details are examples of image-enhancement operations.
- Image enhancement can be performed either in the spatial domain or in the frequency domain. The selection of a particular image-enhancement technique depends on application.
- The spatial-domain image-enhancement technique includes gray-level transformations like image negative, image slicing, image thresholding, gamma correction, etc.
- In a point operation, the value of each pixel is recalculated according to a certain transformation. Point operations commonly involve the adjustment of brightness and contrast in an image.
- The histogram of an image gives the frequency of occurrence of the gray levels. Two different images can have the same histogram. A histogram is invariant to rotation.
- The goal in histogram equalisation is to approximate the grayscale value distribution of an image to the uniform distribution. Histogram equalisation spreads the grayscale values and allows one to see a larger range of grayscale values.
- Spatial domain filtering is accomplished by convolving the image with a filter kernel. Low pass filtering is used to remove the noise at the expense of blurring of the image. High-pass filtering enhances the details in the image.
- Frequency domain filtering is accomplished by the multiplication of the image spectrum by the Fourier transform of the filter kernel. The filtered image in spatial domain is obtained by taking inverse Fourier Transform.
- Gaussian filters are local smoothing operators whose impulse response has the shape of a 2D Gaussian distribution. Gaussian filters can be used to filter out noise in images. The user can vary the standard deviation of the Gaussian and the convolution mask size.
- Image subtraction is useful for background removal. Image multiplication is useful for masking out a portion of an image.

Summary



- The objective of image restoration is to compensate for defects which degrade an image. Degradations are due to motion blur, noise and camera misfocus.
- The deterministic method of image restoration can be effective if information about the degradation is known.
- The general model of image degradation is given by

$$g(m, n) = f(m, n) * h(m, n) + \eta(m, n)$$

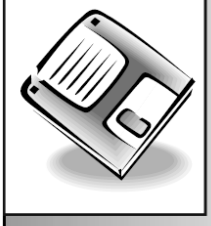
where $f(m, n)$ represents the input image, $h(m, n)$ represents the point spread function and $\eta(m, n)$ represents the noise in the image.

- The inverse filter is an effective method to minimise the degradation in the absence of noise. The pseudo-inverse filter is an alternative to inverse filter if inverse does not exist.
- The wiener filter is an optimal filter with respect to minimum mean square error. It will reduce to an inverse filter in the absence of noise.
- Blind deconvolution is an effective method to restore an image if information about the degradation is not known.
- The different blind-deconvolution techniques include priori blur identification, non-parametric finite support identification, ARMA parameter estimation and multichannel blind deconvolution.
- A median filter is basically a non-linear filter. It is effective in minimising salt-and-pepper noise.
- The performance metrics related to image restoration are (i) Blind Signal to Noise Ratio (BSNR), and (ii) Improvement in Signal to Noise Ratio (ISNR).

Summary

- The objective of image segmentation is to partition an image into meaningful parts that are relatively homogenous in a certain sense.
- Local segmentation deals with segmenting sub-images which are small windows on a whole image whereas global segmentation is concerned with segmenting a whole image.
- Image segmentation can be performed in three different approaches which are (i) region approach, (ii) boundary approach, and (iii) edge approach.
- Region growing is an approach to image segmentation in which neighbouring pixels are examined and added to a region class if no edges are detected. Region splitting is a top-down approach and it begins with a whole image and divides it up such that the segregated parts are more homogenous than the whole.
- The clustering technique attempts to access the relationships among patterns of the data set by organising the patterns into groups or clusters such that patterns within a cluster are more similar to each other than patterns belonging to different clusters.
- An image can be segmented by the thresholding operation in which a histogram of the image can be used as a tool to select the threshold value.
- Edges are basically discontinuities in the image intensity due to changes in the image structure. Robert, Prewitt, Sobel and Frei–Chen kernels can be used to detect the edges in an image.
- The Canny edge-detection algorithm is the most powerful and widely used algorithm used to detect edges in an image. The three phases in the Canny edge-detection algorithm are (i) smoothing and differentiation, (ii) non-maximal suppression, and (iii) thresholding.
- The snake algorithm and the watershed algorithm are powerful algorithms used to segment an image. In the watershed transformation, the image is considered as a topographic surface.
- Chain codes and polygonal approximation are widely used to represent the shapes in an image.

Summary



- Patterns are basically regularity in the input data. Patterns can be captured either in geometric or statistical terms. Object recognition deals with the recognition of the patterns in an image.
- Patterns can be represented in one of the following forms:

  (a) vector form, (b) string form, and (c) tree representation
- Different approaches to pattern recognition include

  (a) statistical approach, (b) structural approach, and (c) neural network approach
- The bayesian approach is a popular statistical approach to object recognition. In statistical approach each pattern is represented as a measurement and is viewed as a point in the *d*-dimensional space. Given a set of training patterns from each class, the objective is to establish decision boundaries in the feature space which separates patterns belonging to different classes.
- One of the simplest approaches to pattern recognition is template matching. In template matching, the correlation between the input image and target is taken into account.
- In the neural-network approach, if the target is known then it comes under the category of supervised learning, and if the target is unknown then it comes under the category of unsupervised learning.
- The multi-layer perceptron is found to be more effective in object recognition than the single-layer perceptron.
- In structural approach, the structural features, commonly known as primitives, are assumed to be known. Structural pattern recognition assumes that pattern structure is quantifiable and extractable so that structural similarity of patterns can be accessed.
- Pattern recognition finds application in different areas like industrial automation, document image analysis, biometric recognition, and remote sensing.

Summary

- Compression is compact representation. Image compression deals with the reprsenation of an image with minimum number of bits.
- Compression is achieved by minimising the redundancies in an image. Different types of redundancies in an image are (i) spatial redundancy, (ii) temporal redundancy, and (iii) psychovisual redundancy.
- Two broad classifications of image compression are (i) lossless compression, and (ii) lossy compression. In lossless compression, the reconstructed image exactly resembles the original image. In lossy compression, there is loss of information.
- The following techniques are common in lossless image compression: (a) run-length coding, (ii) Huffman coding, (iii) arithmetic coding, (iv) Shannon–Fano coding, and (vi) Dictionary-based techniques like LZW.
- Quantisation is basically approximation. Quantisation is a non-linear and irreversible process. Quantisation can be broadly classified into (a) scalar quantisation, and (b) vector quantisation.
- Different types of vector quantisation schemes include (i) mean removed vector quantisation, (ii) gain-shape vector quantisation, (iii) multistage vector quantisation, (iv) tree-structure vector quantisation, (v) classified vector quantisation, and (vi) hierarchical vector quantisation.
- The different types of predictive coding techniques are (i) delta coding, (ii) DPCM, and (iii) ADPCM techniques.
- The transform which is widely used in JPEG compression scheme is the Discrete Cosine Transform (DCT). DCT-based image compression is basically a block-based approach and it suffers from blocking artifacts.
- Wavelet-based image-compression scheme overcomes the problem of blocking artifact. Wavelet transform is employed in the JPEG2000 standard.

Summary

- Morphological image processing is based on the idea of probing an image with a small shape or template known as a structuring element. It is possible to use structuring elements of different shapes to perform a specific task.
- The basic morphological operations are dilation and erosion.
- In the dilation operation, the image is probed with the structuring element by successively placing the origin of the structuring element to all possible pixel locations; the output is 1 if the structuring element and the image have a non-zero intersection and 0 otherwise.
- In the erosion operation, the image is probed with the structuring element by successively placing the origin of the structuring element to all possible pixel locations; the output is 1 if the structuring element is completely contained in the image and 0 otherwise.
- The opening operation is defined as erosion followed by dilation. The opening operation has the effect of eliminating small and thin objects, smoothing the boundaries of large objects without changing the general appearance.
- The closing operation is defined as dilation followed by erosion. The closing operation has the effect of filling small and thin holes in an object, connecting nearby objects and generally smoothing the boundaries of large objects without changing the general appearance.
- Morphological operations such as opening and closing can be regarded as morphological filters that remove undesired features from an image.
- The hit-or-miss operation probes the inside and outside of objects at the same time, using two separate structuring elements. A pixel belonging to an object is preserved by the hit-or-miss operation if the first structuring element translated to that pixel fits the object, and the second translation element misses the object. The hit-or-miss operation is useful for the detection of specific shapes.
- The thinning operation accomplishes erosion without breaking objects. Thinning can be accomplished as a two-step approach, with the first step being erosion that marks all candidates for removal without actually removing them; and in the second pass, candidates that do not destroy the connectivity are removed.

Summary

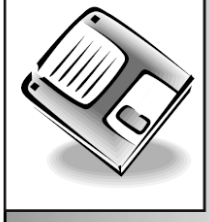
- Visible light is in the range 400 nm (blue) to 700 nm (red).
- Cones have three different kinds of colour-sensitive pigments, each responding to a different range of wavelengths.
- The combination of the responses of these different receptors gives us colour perception which is generally termed the tristimulus model of colour vision.
- CIE uses three primaries X, Y and Z to replace red, green and blue.
- The Y component was chosen to correspond to luminous efficiency

  Y= luminance X, Z = chromaticity
- The normalised values are given by

$$x = \frac{X}{X+Y+Z}; y = \frac{Y}{X+Y+Z}; z = \frac{Z}{X+Y+Z}$$

- The colour space spanned by a set of primary colours is called a colour gamut.
- The different colour models are RGB, CMY, YC_bC_r, YIQ, etc.
- The RGB model is used most often for additive models. In an RGB model, adding a colour and its complement create white.
- The most common subtractive model is the CMY model. In a CMY model, adding a subtractive colour and its complement create black.
- Each pixel of the colour image will have three values, one each for the red, green and blue components.
- The histogram of a colour image can be considered to consist of three separate one-dimensional histograms, one for each tricolour component.

Summary



- A wavelet is an oscillatory functions of finite duration. Its spectrum resembles the transfer function of a bandpass filter.
- A set of basis functions for a wavelet transform can be generated from dilations and translations of a basic wavelet.
- The continuous wavelet transform represents a signal as a function of time and scale.
- The discrete wavelet transform represents an N-point signal with N coefficients. It represents an N-by-N image with N^2 coefficients.
- Discrete Wavelet Transform (DWT) can be implemented through (a) filter bank, or (b) lifting scheme.
- Multi-resolution analysis is possible with wavelet transform. Through multi-resolution analysis, a signal can be analysed at different scales.
- Wavelets are widely used in the field of image compression, image denoising and image watermarking. The choice of the wavelet depends on the nature of the input image as well as the nature of application.
- Progressive image coding is possible through wavelet transform. The most commonly used technique to encode wavelet coefficients are (i) EZW, (ii) SPIHT, (iii) SPECK, and (iv) EBCOT.
- The different types of wavelet-thresholding methods for image denoising are (i) VisuShrink, (ii) LevelShrink, (iii) SureShrink, and (iv) BayesShrink.