|  |  |
| --- | --- |
| Video | Description |
| 1 | * When a human sees an image, he usually notices   + Objects   + Properties of objects   + Interaction of objects   + Scene setting   We can describe the image with these four |
| 2 | * According to research, information relies on two levels   + Syntax: information present directly in the image, no prior human knowledge required. It is limited.   + Semantics: information that is present in the image and also requires prior human knowledge, experience. It is unlimited |
| 3 | * Content Words: Important words in a sentence. Also called keywords. * Digital Image Processing is the processing of digital images using digital computer algorithms. * In computer vision a camera is linked to a computer. The computer interprets images of a real scene to obtain information useful for tasks such as navigation, manipulation and recognition. * Vision is about discovering from images **what** is present is the image and where it is. * A digital image is a representation of a two-dimensional image as a finite set of digital values, called picture elements or pixels. * Five W’s: What, where, who: identity, when: time, why * Why is related to semantics, rest of the W’s refer to syntax * When an algorithm is to be written, there are three things that are important   + Input   + Output   + Processing |
| 4 | * Pixel is concatenation of one row and one column of an image matrix. A cell in image matrix. * Pixel values typically represent gray levels, colors, heights, opacities etc. * Digitization implies that a digital image is an approximation of a real scene. * Pixels can be of three types   + Grayscale   + Black and white (Binary)   + Color (RGB) * Not able for light to pass through, Not transparent * When an image is converted into digital image then some information is lost because there is difference between analogue and digital signal. Human sees an analog image as compared to computer that sees a digital image. |
| 5 | * Digital image processing focuses on two major tasks   + Improvement of pictorial information or Quality for human interpretation. e.g. Removing blur, restoring color contrast of image taken under the sun.   + Processing of image data for storage, transmission and representation for autonomous machine perception (scope). * Goals of image processing may differ with scope. |
| 6 | * Human sees image in which there are objects, interaction of objects, semantic information. * Computer sees what values are present as digits that may be binary numbers or RGB colors * Computer vision course is about converting scene information into digits and digit information into scene. * Image processing: left to right, how computer will understand what is present in the image. Conversion of human understanding into computer understanding. * Image synthesis: right to left, also called computer graphics, image is generated through computing mechanism and software tools. Conversion of computer understanding into human understanding. |
| 7 | * Computer vision is broader field that has two branches   + Image processing   + Computer graphics * Segmentation: Partitioning the image * DIP: Input is image and output are image. Low level process * CV: High level process. Input is image and output are understanding |
| 8 | * Computers are better at easy things * Humans are better at hard, complex things in which number of objects with their properties is very high |
| 9 |  |
| 10 | * 3d models are made for approximating final output. * OCR, Technology used to convert scanned document to text * Face detection * Smile detection * Object recognition * Iris recognition * Robotics * Medical imaging |
| 11 | Steps in Digital image processing:   * Problem Domain: defines scope * Image acquisition: acquire image using camera * Image enhancement * Image restoration: restore image in its actual size, format and quality * Morphological processing: extract features of an image that are low level features such as shape, texture, color, region, outer boundary * Segmentation * Object recognition * Representation and description: drawing bounding boxes and labeling them   Optional processing stages   * Image compression: to make transportation of image convenient. * Color image processing |
| 12 | * Lens focuses light from objects onto the retina * Retina is covered with light receptors called cones (6-7 million) and rods (75 – 150 million) * Cones are concentrated around the fovea and are very sensitive ti color * Color blindness is due to inability to differentiate among different color |
| 13 | * Human visual system can percieve approximately 10^10 different light intensity levels * Human visual system can discriminate between a much smaller number aka brightness adaptation * Human vision cannot correctly discriminate Mach Bands * Simultaneous contrast refers to the way in which two different colors affect each other in the way we percieve the tone when placed side by side * Optical Illusions are drawback of Human vision |
| 14 | * Light is just a particular part of electromagnetic spectrum that can be sensed by human eye * electromagnetic spectrum is split up according to the wavelengths of different forms of energy the **band** * Image applications include electromagnetic spectrum far from visible spectrum such as x-ray, UV, infrared etc * Colors that are prceived by human eye are determined by the nature of the light reflected from an object * Human visual system can sense light from 380 to 740 nm |
| 15 | * Image acquisition:Images are typically generated by illuminating a scene and absorbing the energy reflected by the objects in that scene. In image processing, it is defined as the action of retrieving an image from some source, usually a hardware-based source for processing * Image representaion: Most of the time, it refers to the way that the conveyed information, such as color, is coded digitally and how the image is stored * Pixel: Concatenation (intersection) of one row and one column in image matrix. It is unit of a digital image. * In image, there are M rows and N columns. * Pixel values are most often grey levels in the range 0-255(black-white) * Application of Illumination and scene are   + X-ray of a skeleton   + Ultrasound of an un born baby   + Electron-microscopic images of molecules |
| **16.** 1:31 | * Human visual system falls in analog domain and computer vision falls in digital domain * A digital sensor can only measure a limited number of samples as a discrete set of energy levels. * Quantization is the process of converting a continous analog signal into a digital represention of this signal * Image sampling: The process of digitizing the co-ordinate values is called Sampling. A continuous image f(x, y) is normally approximated by equally spaced samples arranged in the form of a NxM array where each elements of the array is a discrete quantity. Related terms are   + number of pixels in an image   + spatial digitization   + along x-axis * Image Quantization: A lossy compression technique achieved by compressing a range of values to a single quantum value. Related terms are   + grey levels digitization   + time slicing   + along y-axis * Bins: A clusture of close values. * Digital image is always an approximation of real world scene. Information loss * Upsampling: Upsampling is the process of inserting zero-valued samples between original samples to increase the sampling rate * Downsampling: To make a digital audio signal smaller by lowering its sampling rate * Sample rate: The sampling rate determines the spatial resolution of the digitized image * Quantization Level: It determines the number of grey levels in the digitized image * How much samples are needed:Depends upon the nature of application. * Effect of noise: Sampling is not affected but affects quantization |
| **17**  3:54, 5:30 | * Pixel: Smallest element of an image, can store a value proptional to intensity of light, located at particulat location * No. of pixels: MxN * To see more details while zooming, increase number of pixels * Oversampling: More sampling than needed. Need much more storage to save redundant and useless details * Image zooming: A function of focusing on a section of an image and increasing its overall size for greater detail * Pixel resolution: Total number of pixels in an image. * Size = number of pixels X bit size of each pixel |
| **18**  2:32 | * Spatial Resolution: Smallest descernible detail in an image i.e. number of independent pixel values per inch * The number of pixels in an image does not matter. * Spatial resolution of an image is determined by how sampling was carried out. * Terms:   + Vision specialists: pixel size   + Graphic designers: Dots per inch (DPI)   + Monitors: Dots per inch (DPI)   + Laser printers: Lines per inch (LPI)   + Tablets and Mobile phones: Pixels per inch (PPI) * DPI: How many dots of ink are printed per inch when an image get printed out from the printer. There may be many dots per inch used for printing one pixel. Most of the printers use CMYK (Cyan, Magenta, Yellow, Black) color scheme so colors are limited. Printer has to choose one color for many shades * Printers use 300DPI to 600DPI and more. |
| 19 | * Segmentation: Subdivides an image into its constituent regions or objects. Seperates and image into regions which are called as objects and background. What part of image is of object and what is not. We put labels on image for human and machine perception. We put one on forground and zero on background. * Using segmentation we find continuity of an object and discontinuity between objects * Discontinuity: Partition based on abrupt changes in intensity. E.g. edges in an image/point/line/edge/corner detection * Similarity: Partition based on intensity similarity. E.g. Thresholding , region growing/splitting/merging * Segmentation based on color: Based on greyscale, leads to inaccuracy in object labeling * Segmentation based on texture: Separate the pattern of intensity values. Enables object surfaces with varying patterns of grey to be segmented * Segmentation based on motion: The task of identifying the independently moving objects. Subtract the next image from previous in video to find the moving objects. * Segmentation based on depth: Done using Laser range finder to find the distance of various objects from the sensor * Segmentation based on shape: Based on boundaries of an object. Partitioning a given shape into several components or segments that capture application-specific part–whole relations as well as possible |
| 20 | * How much Sampling is enough:   + Does the image look aesthatically pleasing?   + Can you see what you need to see in the image?   + Depends upon the scope:     - Count Vehicles     - Read number plate * Oversampling increases cost of an image. * Cost of an image: Resource required in terms of Size, Transportation, Filtering and other Processes |
| 21 | Image Enhancement: Make image quality better |
| 22 |  |
| 23 | r is input grey level. C is the constant for controlling the value. |
| 24 |  |
| 25 | * Highlighting interesting details: Improve the quality of individual objects so that their inner details are visible, sharpening the boundaries so that they are differentiatable. * Removing noise from images: Removing unwanted overlaping things, restoring blurred image * Making images more appealing to human eye * There are two broad categories of image enhancement   + Spatial domain techniques   Direct manupulation of image pixels such as point processing   * Frequency domain techniques that is Manipulation of fourier transform or wavelet transform of an image. |
| 26 | * Histogram of an image shows us the distribution of grey levels in the image. It means counting of frequency of something. * Used in segmentation * Histogram equalization: Make grey levels to be equally distributed in the image * High contrast image has most evenly spaced histogram * Histogram is one of the fundamental technique in image enhancement. |
| **27**  2:57 | * Image noise: Unwanted information in an image. Random change of intensity value in image pixels. * Noise arise during image acquisition and transmission * Noise cause image degradation * Image sensors can be affected by ambient conditions * Interference while image transmission cause noise * Quality comparison of image formats:   + GIF<PNG<JPG, JPEG< TIFF |
| 28 | * The purpose of image restoration is to “compensate for” “or undo” defects which degrade an image. Degradation comes in many forms such as motion blur, noise, and camera misfocus. |
| 29 | * Noise Models: Based on statistical distribution of data, η(x, y)   + Gaussian (normal)   + Impulse (salt and pepper)   + Uniform   + Rayleigh   + Gamma (Erlang)   + Exponential |
| 30 | * Derivative and diffrentiation produces noise * Laplacian filter: Based on 2nd order derivative * Not used by itself because it is too sensitive to noise * Usually when used for edge detection the laplacian is combined with a smoothing gaussian filter |
| 31 | * Geometric mean filter is better than arithimatic mean filter because it tends to lose less information * The contra-harmonic mean filter is member of a set of nonlinear mean filters which are better at removing Gaussian type noise and preserving edge features than the arithmetic mean filter |
| **32**  3:24, 5:55 | * Image restoration attempts to restore images that have been degraded   + Identify the degradtion process and attempt to reverse it   + Similar to image enhancement, but more objective   + Need a reference image * Objective of image restoration: To recover a distorted image to the original form based on idealized models * The distortion is due to:   + Image degradation in sensing environment e.g random atmospheric turbulance   + Noisy degradation from sensor noise   + Blurring degradation due to sensors e.g camera motion or out-of-focus   + Geometric distortion e.g earth photos taken by a camera in satellite |
| 33 | * Harmonic mean filter is used to remove salt noise * Geometric mean filter is used to remove pepper noise * Contraharmonic mean is the combination of Geometeric and Harmonic mean |
| 34 | * Spatial filters that are based on ordering the pixel values that make up the neighbourhood operated on by the filter   + Median filter   + Max and Min filter   + Midpoint filter   + Alpha trimmed mean filter * Sources of noise:   + Hardware device limitation   + Ambient conditions   + Image compression   + Image transmission   + Image enhancement   + Personnel fault e.g. Out-of-focus |
| 35 | * Segmentation: Subdivides an image into its constituent regions or objects. Seperates and image into regions which are called as objects and background. What part of image is of object and what is not. We put labels on image for human and machine perception. We put one on forground and zero on background. * Color Segmentation is the easiest * Segmentation is the first step in object recognition * There are Three major steps in ML or DIP   + Pre-processing   + Feature extraction   + Classification |
| **36**  5:08 | * Edge: A set of connected pixels that lie on the boundary between two regions * First order derivative tells us where an edge is * Second order derivative can be used to show the edge direction * Derivative cause noise. Higher the order, larger the noise produced * Some derivative filters are:   + Sobel (2nd)   + Prewitt (2nd)   + Roberts (1st)   + Laplacian (2nd) |
| **37**  0:36,  1:45 | * There are three basic typesf of grey level discontinuities that we tend to look for in digital images:   + Points   + Lines   + Edges * We find discontinuites using masks and correlation |
| 38 | * Morphological image processing: Used to extract image components for representation and description of region shape, such as boundaries, skeletons and the convex hull * Morphological image processing (or morphology) describes a range of image processing techniques that deal with the shape (or morphology) of features in an image * Morphological operations are typically applied to remove imperfections introduced during segmentation, and so typically operate on bi-level images |
| 39 | * Structuring Element: A shape, used to probe or interact with a given image, with the purpose of drawing conclusions on how this shape fits or misses the shapes in the image * Fit: All on pixels in the structuring element cover on pixels in the image * Hit: Any on pixel in the structuring element covers an on pixel in the image * Miss: No pixel in the structuring element covers a pixel in the image |
| 40 | * Fundamentally morphological image processing is very like spatial filtering * The structuring element is moved across every pixel in the original image to give a pixel in a new processed image * The value of this new pixel depends on the operation performed * There are two basic morphological operations: erosion and dilation |
| 41 | * More interesting morphological operations can be performed by performing combinations of erosions and dilations * The most widely used of these compound operations are   + Opening   + Closing |
| 42 | * Using the simple technique we have looked at so far we can begin to consider some more interesting morphological algorithms   + Boundary extraction   + Region filling   + Extraction of connected components   + Thinning/thickening   + Skeletonisation * The purpose of morphological processing is primarily to remove imperfections added during segmentation * Using the basic operations we can perform opening and closing * More advanced morphological operation can then be implemented using combinations of all of these |
| 46 | * Object Recognition is used in   + Photography   + Assisted Driving: Autonomous vehicle |
| 47 | * Template matching is the easiest method for object recognition. * Fit is can be defined by Morphological fit or Correlation * Correlation ranges from -1 to 0 to 1 * Mask or filter is cropped image image that is to be correlated in the image. |
| 48 | * Challenges of object recognition:   + View point variation   + Illumination or Lighting conditions   + Occlusion: Hurdle between camera and the object of interest. E.g. glasses weared while eyes detection   + Scale   + Deformation |
| 49 | * In view point variation challenge, our aim is to train the algorithm at different angles * Modeling variability: Train the algorithm at multiple templates. * There is another challenge of object recognition that is   + Camera position   + Illumination   + Internal parameters   + Within class variation: different chairs |
| 50 | * Timeline of object Recognition   + 1965 to late 1980: alignment, gemetric primitives * Alignment: Fitting a model to a transformation between pairs of features (matches) in two images. Bring something towards its actual angle * Geometric primitives are   + Translation   + Rotation   + Scaling * Position of one object is different from the other. Change the actual location of object within image. * Object recognition is about   + Get different set of images   + Extract features from them   + Improve the results using these features |
| 51 | * To identify object there should be two things   + Color   + Shape * Early 1990s: invariants, appearnce-based methods * Emperical models of image variability: Appearance based techniques. Things that human can understand by viewing. * Color Histograms are used in object recognition. Color that is present in bigger amount will have larger cluster. * There are two types of techniques in image processing   + Global: Processing the image as a whole   + Local: Processing a specific part of image |
| 53 | * Mid to late 1990s: Sliding window approaches * Late 1990s: feature-based methods, Shift of machine learning * Local features: Combining local appearance, spatial constraints, invariants and classification techniques from machine learning. * Feature based methods: object recognition based on local features. Global features are broke down into local features then algorithms are applied |
| 54 | * Early 2000s: Parts and shape based methods Geometry: location of parts in the object * Objects are made up of parts.   + Generative representation: bottom up approach * Model: Relationship among the parts of the object   + Relative locations between parts   + Appearance of part * Issues:   + How to model location   + How to represent appearance   + Sparse or dense (pixels or regions)   + How to handle occulsion/clutter(mixing) |
| 55 | * 2003 to present: bag of features * Work on object features rather than geometry * Texture: repitition of pattern * Produce generative models using bag of features |
| 56 | * Present trends: combination of local and global methods, modeling context, integration recognition and segmentation * Global models include ‘gist’ of a scene * Data augmentation: Generation of new data |
| 57 | * Top down approach: Starting from label or starting from global towards local * Bottom up: Local to global, Generative or Discriminative * Posterior probability: probability of something after occurance of an event * Descriminative or global or top down methods model posterior * Generative or feature based methods model likelihood and prior |
| 58 | * Issues in object recognition   + Representation: How to represent an object category. What features should be important, appearance based, color based.     - Generative/Descriminative/hybrid     - Appearance only or location and appearance.     - Invariances       * View point       * Illumination       * Occulsion       * Scale       * Deformation       * Clutter     - Part based or global window/sub-window     - Use set of features or each pixel in image   + Learning: How to form the classifier, given training data     - Variability in training data.     - Unclear how to model categories, so we learn what distinguishes them rather than manually specify the difference – hence current interest in machine learning     - Methods of training: generative vs discriminative     - Level of supervision       * Manual segmentation, bounding boxes, image labels, noisy labels     - Training images:       * Issue of overfitting       * Negative images for discriminative methods Priors   + Recognition: How the classifier is to be used on novel or unknown data. |
| 59 | * Histogram: Frequency of individual grey level inside the image * Histogram Equalization: Histogram Equalization is a computer image processing technique used to improve contrast in images . It accomplishes this by effectively spreading out the most frequent intensity values, i.e. stretching out the intensity range of the image * Image obtained as a result after histogram equalization is known as normalized or equally distributed image w.r.t grey levels |
| 60 | * Neighbourhood operations simply operate on a larger neighbourhood of pixels than point operations * Neighbourhoods are mostly a rectangle around a central pixel * Any size rectangle and any shape filter are possible * Immediate neighbours: pixels located in the ring enclosing the central pixel * Far neighbours: other than immediate neighbours * Point processing: Only a mapping function. No neighbours included in calculation * Image enhancement operations can also work with the values of the image pixels in the neighborhood and the corresponding values of a subimage that has the same dimensions as the neighborhood. * The values in a filter subimage are referred to as coefficients, rather than pixels. * For linear spatial filtering, the response is given by a sum of products of the filter coefficients and the corresponding image pixels in the area spanned by the filter mask * Different terminologies:   + Filter is collection of kernels and hence not unique. It contains coefficients as value.   + Mask usually contain binary values that tells about what part of the image is to hide or show. It contains coefficients as value.   + A window is necessarily a rectangular array of pixels. In most cases, values are a portion of an image. It contains pixel intensity as value.   + Template is not necessarily a rectangle. In most cases, values are a portion of an image. It contains pixel intensity as value.   + Kernel is unique and usually smaller than filter. It contains coefficients as value. |
| 61 | * Objective of neighbourhood operation is to change the value of central pixel based on its neighbours * Median filter is usually better than Min and Max operators * Some neighbourhood operations:   + Min: Set the pixel value to the minimum in the neighbourhood   + Max: Set the pixel value to the maximum in the neighbourhood   + Median: The median value of a set of numbers is the midpoint value in that set |
| 62 |  |
| 63 | Gross details: Overall details of an image  Finer details: Details of individual objects |
| 64 |  |
| 65 |  |
| 66 | * Dealing with Edges in filter application:   + Omit missing pixels (Do not consider missing pixels. Require extra code)   + Pad the image (Famous technique. Add black or white pixels at the edge)   + Replicate border pixels (copy the edge pixels)   + Truncate the image (do not include pixels in resulting image)   + Allow pixels wrap around the image (Copy pixels of opposite edge. Tile effect) |
| 67 | * Image Sharpening: Make individual details prominent. Highlight or enhance details in images   + Sharpening spatial filters seek to highlight fine detail   + Remove blurring from images   + Highlight edges * Procedure   + Sharpening filters are based on spatial differentiation. * Some Applications   + Photo enhancement   + Medical image visualization   + Industrial defect detection |

Q: Why central element is 2^n in laplacian?