CVIT Workshop Session 1

• • •

Image Processing

- Vansh Garg and Akshat Sanghvi

What is an image?

- "A representation of the external form of a person or thing in art."
- It is just a 3D array of numbers, each in the range 0 - 255 (8 bit).
- If the image is of height H and width W, the image will be of dimensions H x W x 3
- Each of these H x W values are called pixels, and are arrays of 3 values, Red , Blue and Green Channels.
- Pixel: It's the smallest unit of an image, representing a single point of light intensity or color

256 shades of grey (Grayscale images)

- A grayscale image uses only shades of gray, ranging from pure black (value of 0) to pure white (value of 255), to represent the variations in light and dark, and represents the intensity or luminance of the pixel.
- To convert a coloured image to grayscale, we just average out the pixel values, i.e., values of Red, Green and Blue channels.

shutterstock.com · 593416619

Different Color Spaces

- Apart from grayscale and normal RGB colors, there is also HSV color space.
- HSV stands for Hue, Saturation and Value.
- Hue: This represents the actual color itself, like red, green, or blue, visualized as an angle on a color wheel
- Saturation: This describes the intensity or purity of the color, ranging from 0 (gray) to 1 (fully saturated color).
- Value: This represents the brightness of the color, going from 0 (black) to 1 (white).

Different Color Spaces
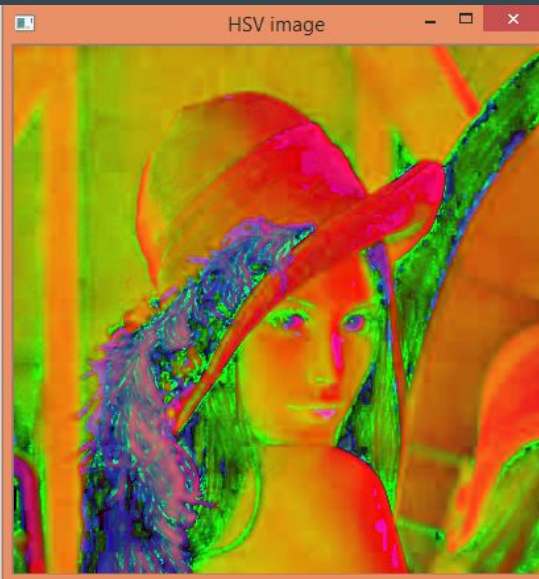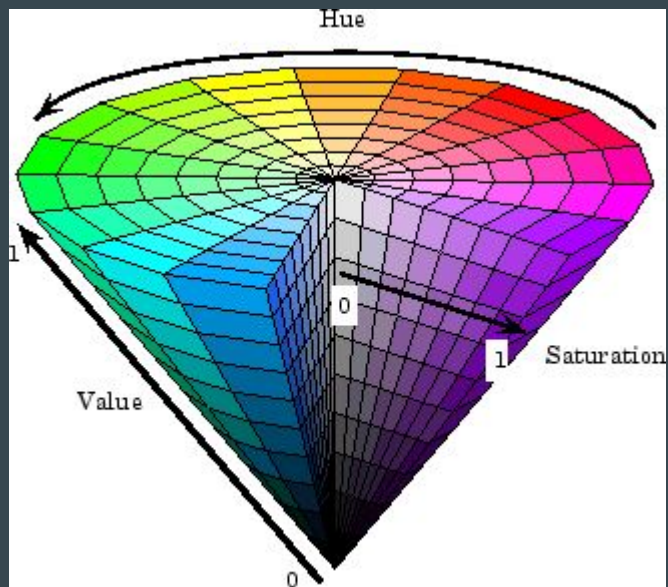
Color Histograms

- Histogram of the R,G,B values. It just shows the color distribution in an image, how much of R,G,B is globally present. It can be used in various applications like contrast stretching, thresholding and noise reduction
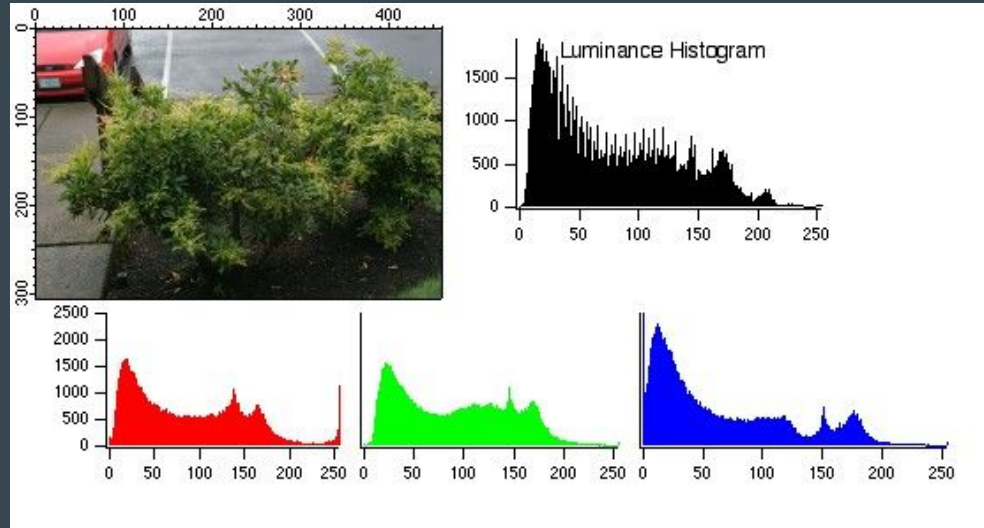
Chroma Keying

- It is the technology behind green screens. It is a special effect where a specific color (usually green or blue) is removed from a video clip
- This creates a transparent / empty area that can be replaced with another image or video, allowing you to place your subject in a completely different background!!!
- We will be trying to implement chroma keying

Emboss

- Embossing an image means to make it appear risen, giving it a sort of 3D effect.
- It is thus mainly used in logos, posters, photography, etc to create a 3D view of the object.

Fourier Analysis

This is a very important topic to study in computer vision and has a wide variety of applications in various algorithms.

Do go through it during the summers in detail, there are many videos explaining it nicely (including 3b1b).

Dilation and Erosion

- Dilation: At each pixel, take the maximum value among neighbouring pixels. This results in adding a layer of 8 white pixels around each white pixel. This is as white pixels have value 255, and black have a value of 0.
- Erosion: Take the minimum of neighbouring pixels, so this results in removing one outer layer of white pixels, essentially
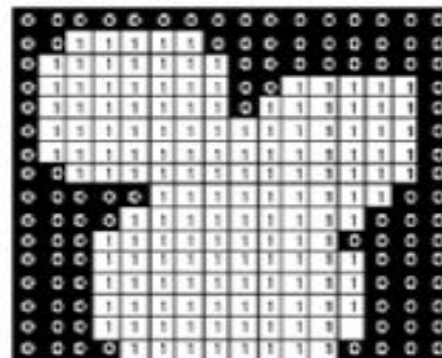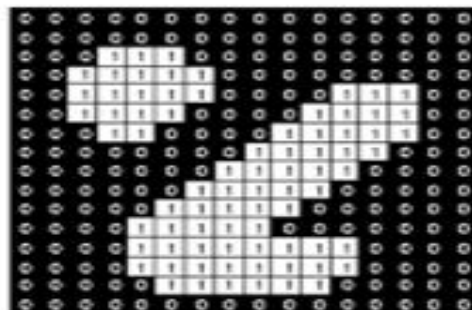
Dilation Operation

Image Pre-processing techniques

Erosion Operation

Fixing Broken Characters

Morphological Opening and Closing

A

$A \ominus B$

$A \circ B = (A \ominus B) \oplus B$

$A \oplus B$

$A \cdot B = (A \oplus B) \ominus B$
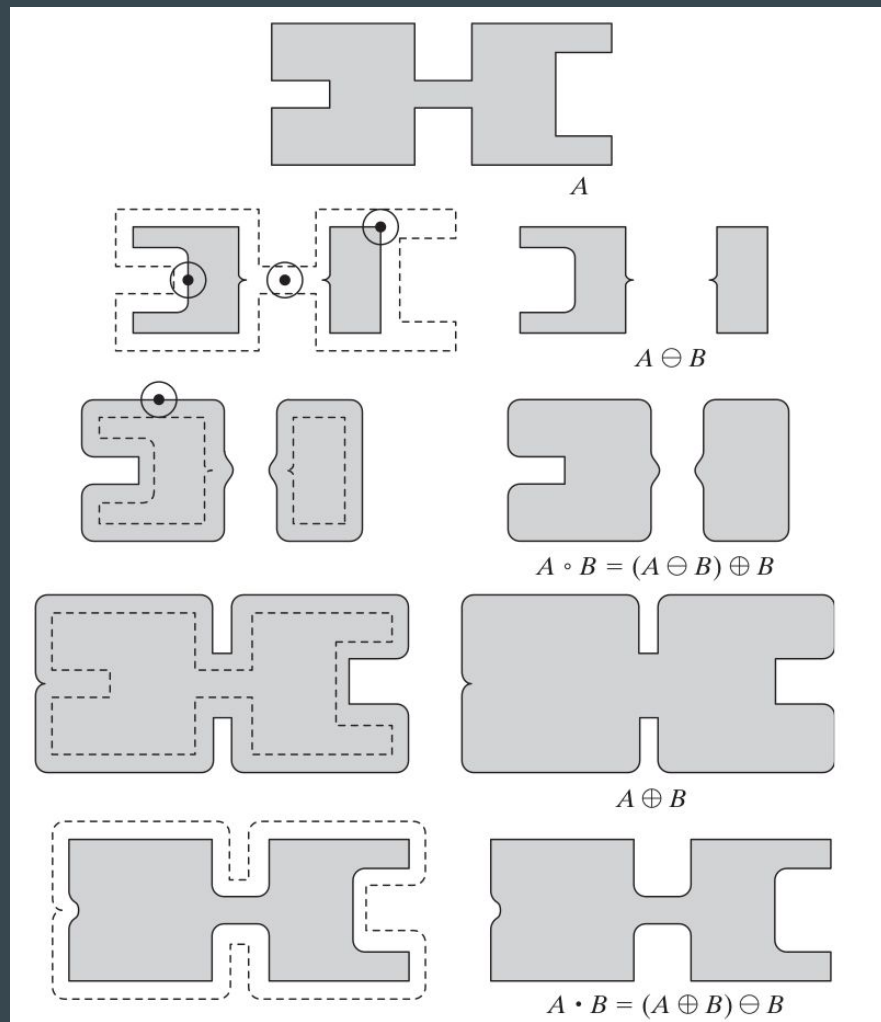
Note

The dilation / erosion kernel size can be more than 1 also. This will involve looking at the "kernel size" amount of neighbourhood, basically more the kernel size, it will involve taking the max / min among more number of pixels, therefore increasing the effect both in the case of dilation and erosion

Closing and Opening Operations

Closing - This operation involves dilation followed by erosion

Opening - This operation involves erosion followed by a dilation

Closing can connect white structures, whereas Opening mainly disconnects them

Closing operation
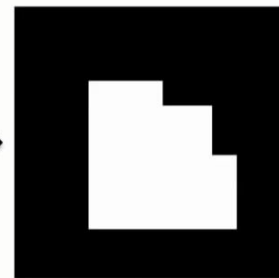
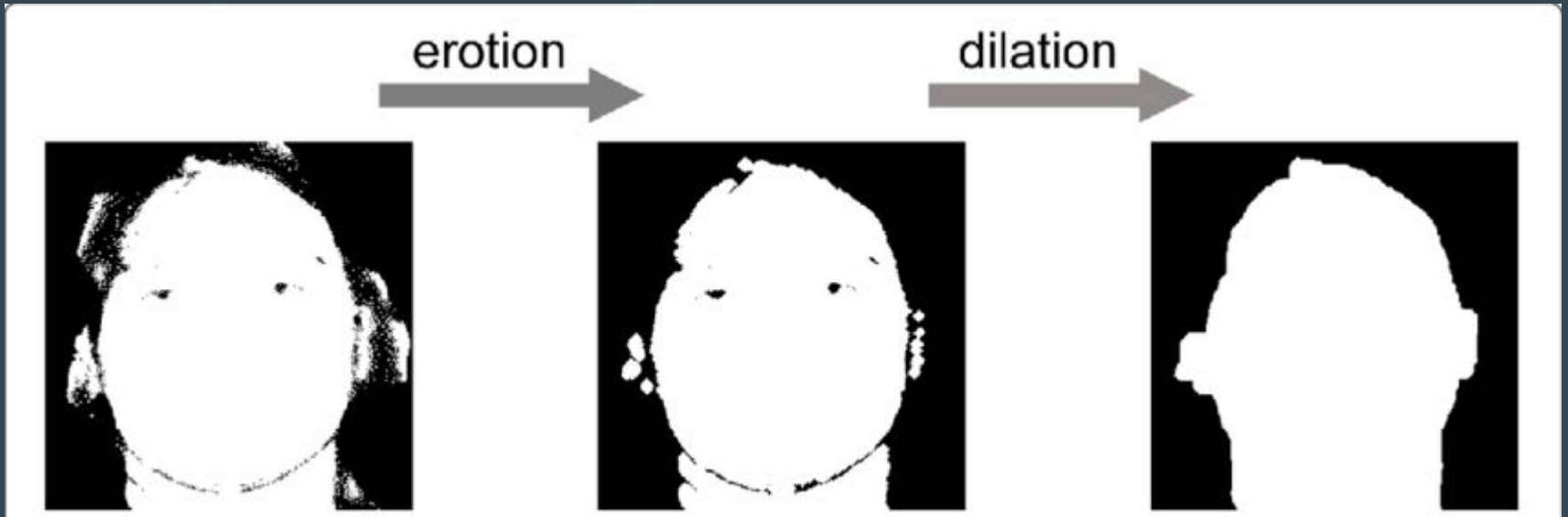Closes the black hole in white components in the image.

Opening operation

Helps in removing the non-major / very small white components

Opening Example