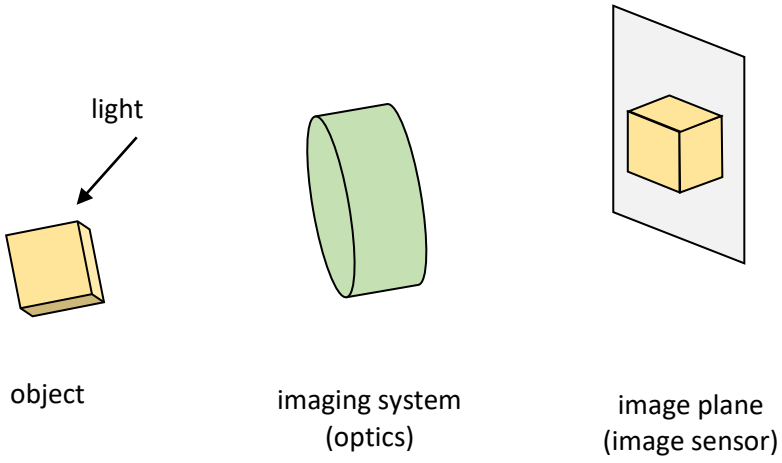
Image Representation

Image formation

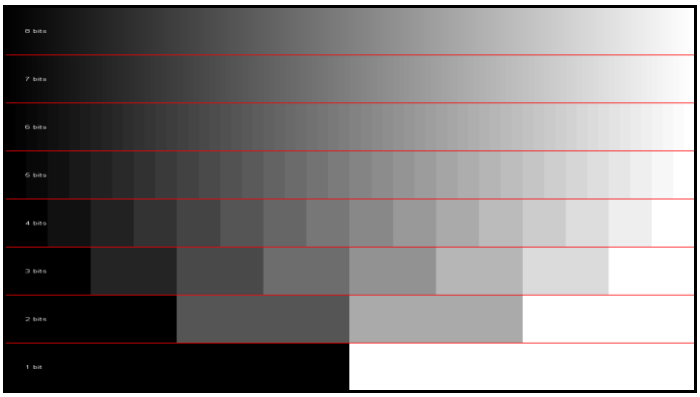


The image on the image plane is formed as a result of the reflected light from the object passing through the lenses (optics). An image sensor is placed on the image plane which consists of an array of small light sensors. The sensor's intensity values are converted into a sequence of digital values and stored as images.

Bitmaps (or raster images) are essentially a collection of dots called **pixels,** the values of which have been derived from the raw sensor values. Pixels are usually square and are arranged in a rectangular grid. The word bitmap comes from an "array of bits that are spatially mapped."

A raster image can be represented with black and white (binary), grayscale or color pixels depending on the capture device, the encoding format and display medium.

Bit depth (or **Color depth)** of an image representation signifies how many bits are used per pixel. Quantization levels depend on the bit depth. Typically, a true color image will have three color channels (Red, Green and Blue) with 8 bits per channel but higher bit depths are also possible. A grayscale image, on the other hand, will only have one channel that represents the intensities of the color-neutral pixels. An example of different levels bit depth (from 1 bit to 8 bits) on a grayscale gradient is shown below.



A **digital camera** will capture an image at or below its maximum resolution (e.g. Nikon - D800 36.3-Megapixel DSLR Camera has 7360 x 4912 pixels.) The number of pixels can be reduced through downscaling for images to take up less space.

File size of an image depends on the maximum resolution of the capture device, its resolution settings, whether it is compressed or not, and its compression method.

File Formats: One needs to differentiate between compressed and raw formats. Compressed formats are lossy but raw formats can lead to large file sizes. A compromise needs to be made keeping quality and storage size in mind.

Common raster image formats include:

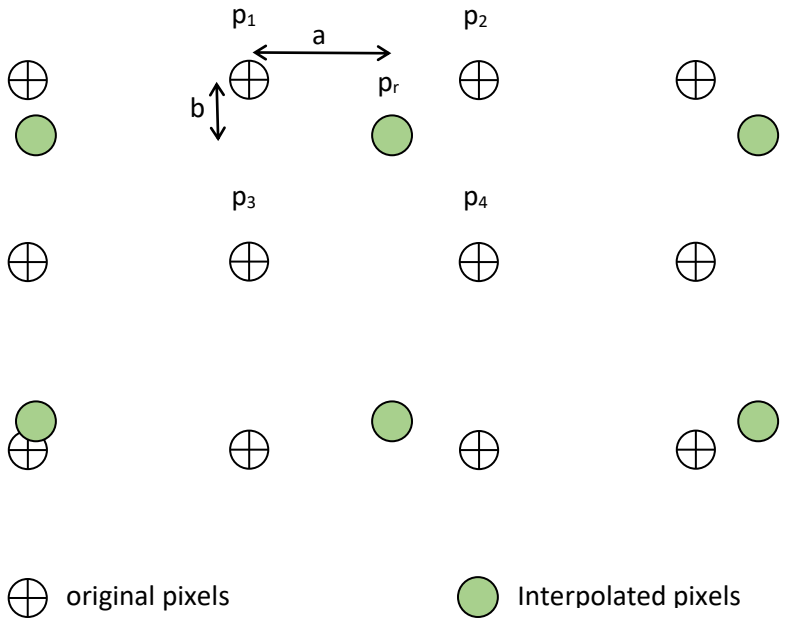
- **BMP** (Windows Bitmap - uncompressed)
- **TIFF** (**TIF**; Tagged Image File Format - 8 or 16 bits per color; can be lossy or lossless)
- **JPEG** (Joint Photographics Expert Group – 24 bit color – lossy compression)
- **GIF** (Graphics Interchange Format - 8-bit palette, or 256 colors)
- **PNG** (Portable Network Graphic - open-source successor to GIF; 8 bit paletted images and 24 bit true color or 48 bit true color with and without alpha channel - lossless)
- **PSD** (Adobe PhotoShop – proprietary format with layers - not a distribution format).

Changing the resolution of an image

To change the number of pixels of a particular image (without cropping) we need to resample the original image at possibly non-integer multiples of the original samples. Therefore, some form of interpolation is required to preserve as much of the information in original image as possible. Different methods can be applied to scale the image:

- Closest pixel: choose the closest pixel that belongs to the original image and use its intensity value as the new pixel's value.

- Bi-linear interpolation: interpolate linearly using the four closest pixels.



$$p_r = (1-a)(1-b)p_1 + a(1-b)p_2 + (1-a)bp_3 + abp_4$$

- where a and b are fractional parts of the x and y coordinates respectively from the top left pixel.
- p_1 - p_4 are the intensity values of the pixels in the original image.
- p_r is the intensity value of the pixel in the resultant image.
- Assume distance between original pixels is 1 in both directions.

- Bi-cubic interpolation: uses curves for interpolation in contrast to bi-linear which uses linear blends between the closest points.

Other operations on images

- Translation
- Rotation
- Skewing
- Cropping
- Stretching along a particular direction