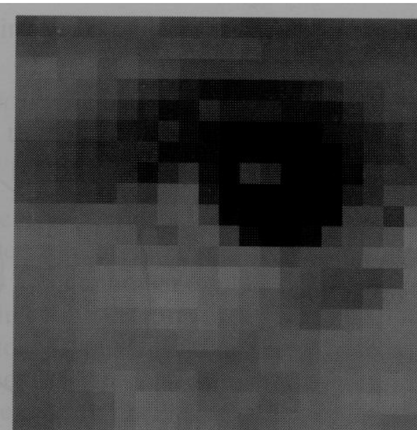
Digital Image Representation

How are images represented in the computer?



117	125	133	127	130	130	133	121	116	115	100	91	93	94	99	103	112	105	109	106
134	133	138	138	132	134	130	133	128	123	121	113	106	102	99	106	113	109	109	113
146	147	138	140	125	134	124	115	102	96	93	94	99	96	99	100	103	110	109	110
144	141	136	130	120	108	88	74	53	37	31	37	35	39	53	79	93	100	109	116
139	136	129	119	102	85	58	31	41	77	51	53	53	33	37	41	69	94	105	108
132	127	117	102	87	57	49	77	42	28	17	15	13	13	17	41	53	69	88	100
124	120	108	94	72	74	72	31	35	31	15	13	15	11	15	13	46	75	83	96
125	115	102	93	88	82	42	79	113	41	19	100	82	11	11	17	31	91	99	100
124	116	109	99	91	113	99	140	144	57	20	20	15	11	15	17	63	87	119	124
136	133	133	135	138	133	132	144	150	120	24	17	15	15	17	20	115	113	88	150
158	157	157	154	149	145	133	127	146	150	116	35	20	19	28	105	124	128	141	171
155	154	156	155	146	155	154	154	147	139	148	150	138	120	128	129	130	151	156	165
150	151	154	162	166	167	169	174	172	167	177	166	164	140	134	120	121	120	127	172
145	149	151	157	165	169	173	179	176	166	166	157	145	136	129	124	120	136	163	168
144	148	153	160	159	158	165	172	165	169	157	151	149	141	130	140	151	162	169	167
144	141	147	155	154	149	156	151	157	157	151	144	147	147	149	159	158	159	166	165
139	140	140	150	153	151	150	146	140	139	138	140	145	151	149	156	156	162	162	161
136	134	138	146	156	164	153	146	145	136	139	139	140	141	149	157	159	161	169	166
136	133	136	135	144	159	168	159	151	142	141	145	139	146	153	156	164	167	172	168
133	129	140	142	146	159	167	165	154	151	146	141	147	154	156	160	161	157	153	154

- A signal is a *function* depending on some variables with physical meaning.
- Signals can be
  - one-dimensional (e.g., dependent on time),
  - two-dimensional (e.g., images dependent on two co-ordinates in a plane),
  - three-dimensional (e.g., describing an object in space),
  - or higher-dimensional.

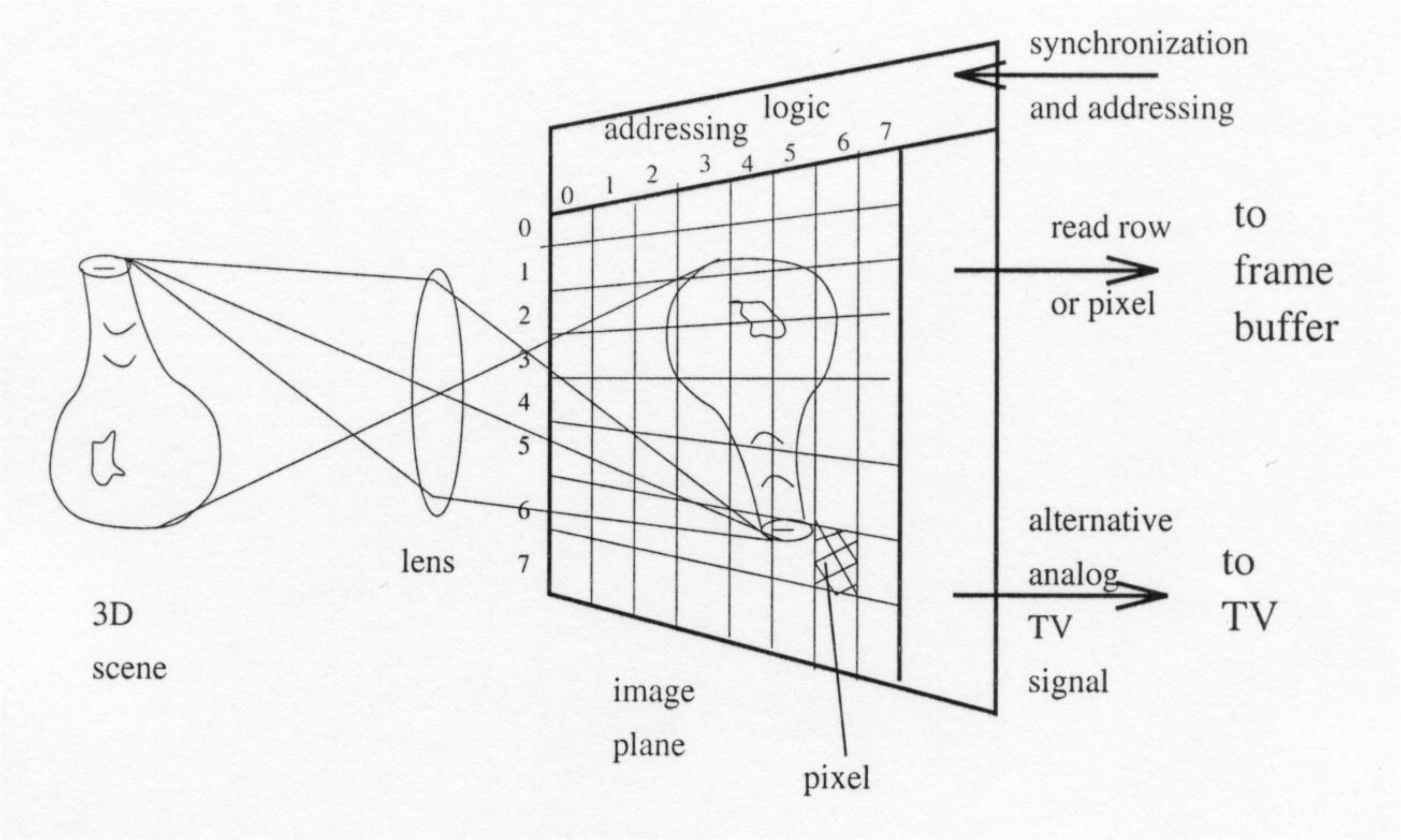
- A scalar function may be sufficient to describe a monochromatic (單色的) image, while vector functions are to represent, for example, color images consisting of three component colors.

Image formation

- There are two parts to the image formation process:
  - The ***geometry of image formation***, which determines where in the image plane the projection of a point in the scene will be located.
  - The ***physics of light***, which determines the brightness of a point in the image plane as a function of illumination and surface properties.

CCD (Charged-Coupled Device)(電荷耦合器) cameras

- Tiny solid state cells **convert light energy into electrical charge**.
- The image plane acts as a digital memory that can be read row by row by a computer.


synchronization
and addressing
addressing logic
0 1 2 3 4 5 6 7
0
1
2
3
4
5
6
7
read row
or pixel
to
frame
buffer
alternative
analog
TV
signal
to
TV
3D
scene
lens
image
plane
pixel

- The ***image*** can be modeled by a continuous function of two or three variables;
- arguments are co-ordinates ***x, y*** in a plane, while if images change in time a third variable ***t*** might be added.
- The image function values correspond to the brightness at image points.

- The image on the human eye retina or on a TV camera sensor is intrinsically (本質上) 2D.
- We shall call such a 2D image bearing information about brightness points an ***intensity image***

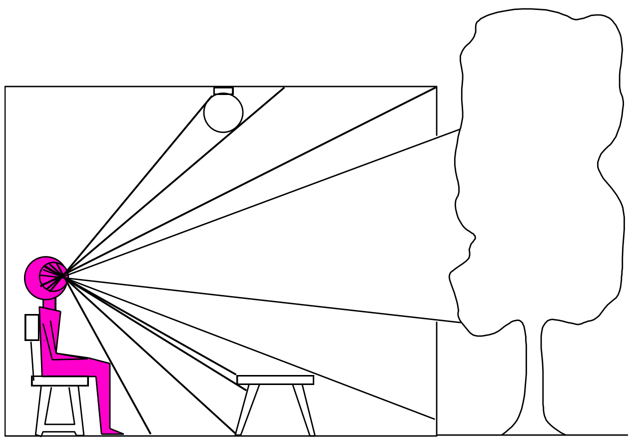
- The real world which surrounds us is intrinsically 3D.
- The 2D intensity image is the result of a perspective projection (透視投影) of the 3D scene.

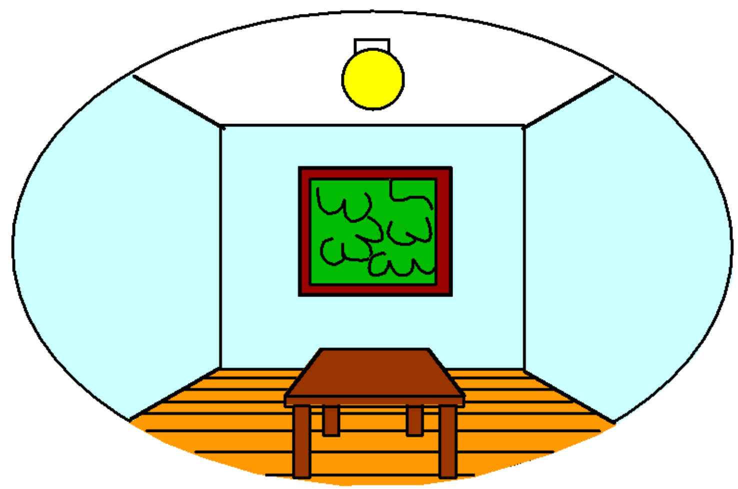
- When 3D objects are mapped into the camera plane by perspective projection a lot of information disappears as such a transformation is not one-to-one.
- Recognizing or reconstructing objects in a 3D scene from one image is an ill-posed problem.

Dimensionality Reduction Machine (3D to 2D)

3D world

2D image



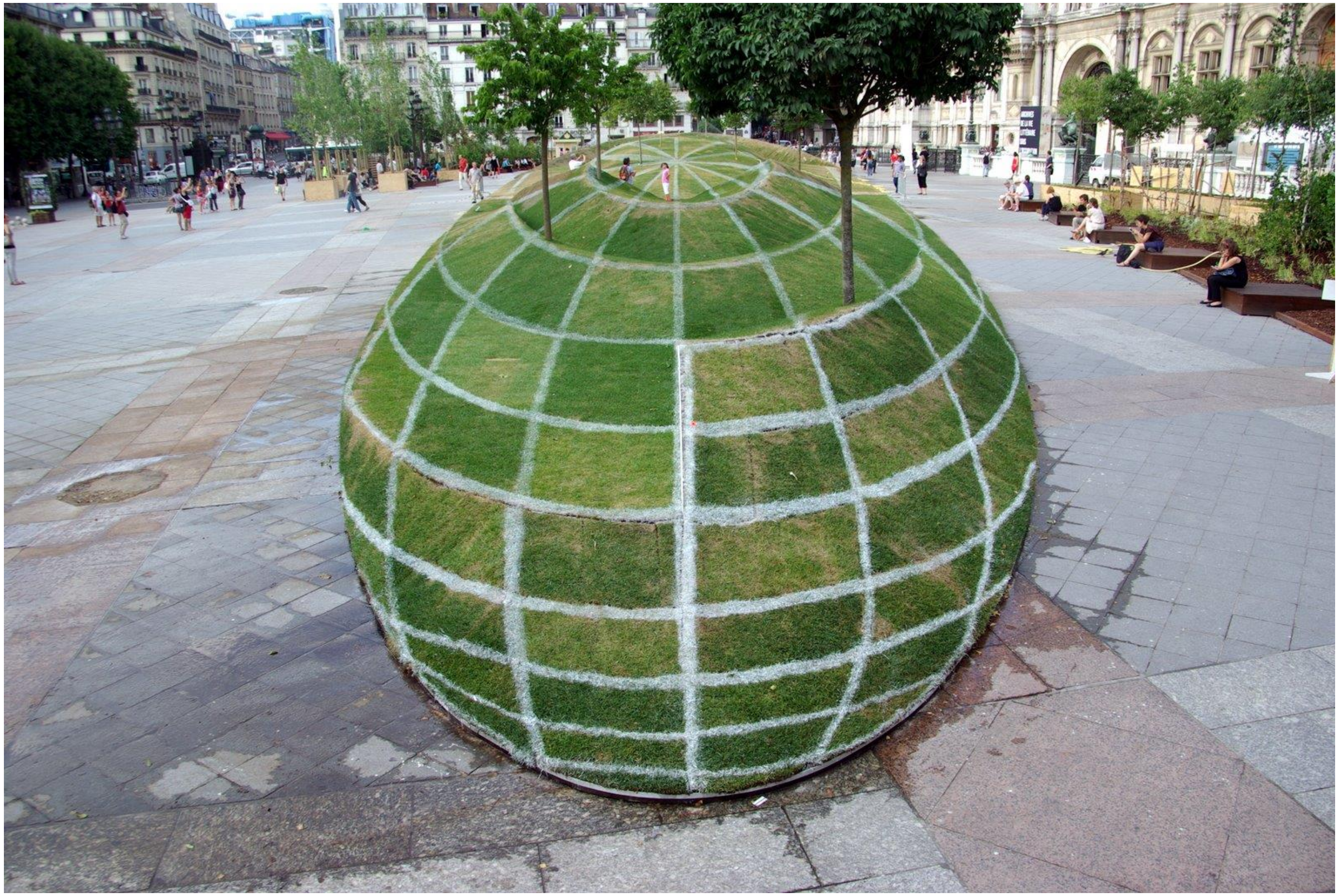


Point of observation

Projection can be tricky…

Projection can be tricky…

Projection can be tricky…

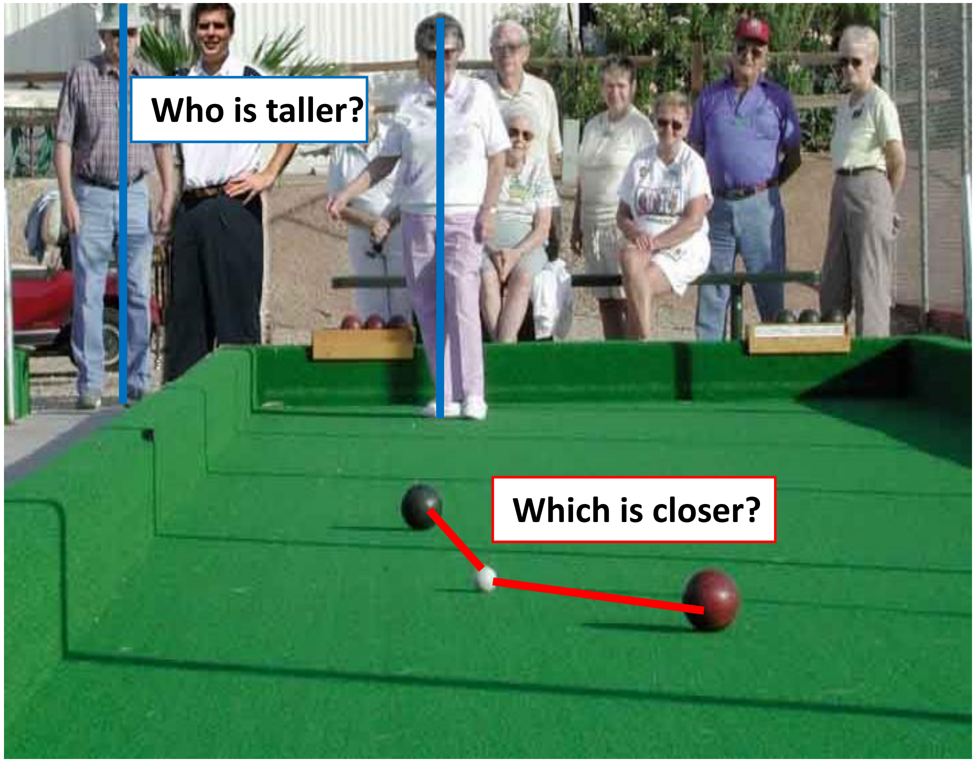


Projection can be tricky…

Projective Geometry

What is lost?

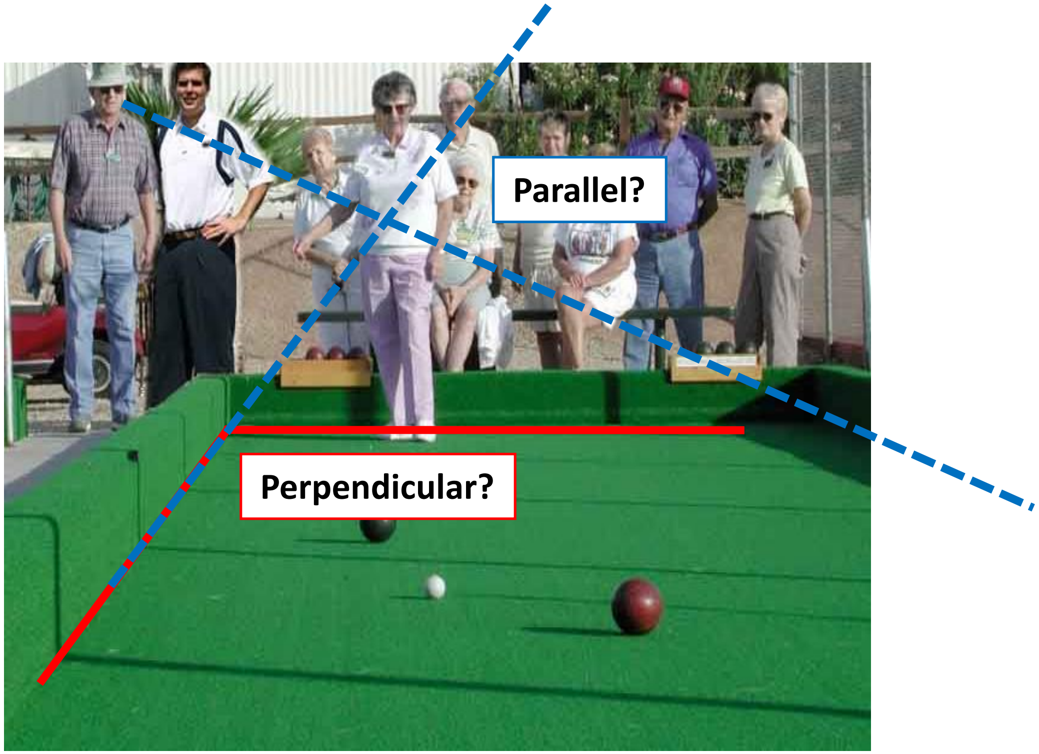
- Length



Projective Geometry

What is lost?

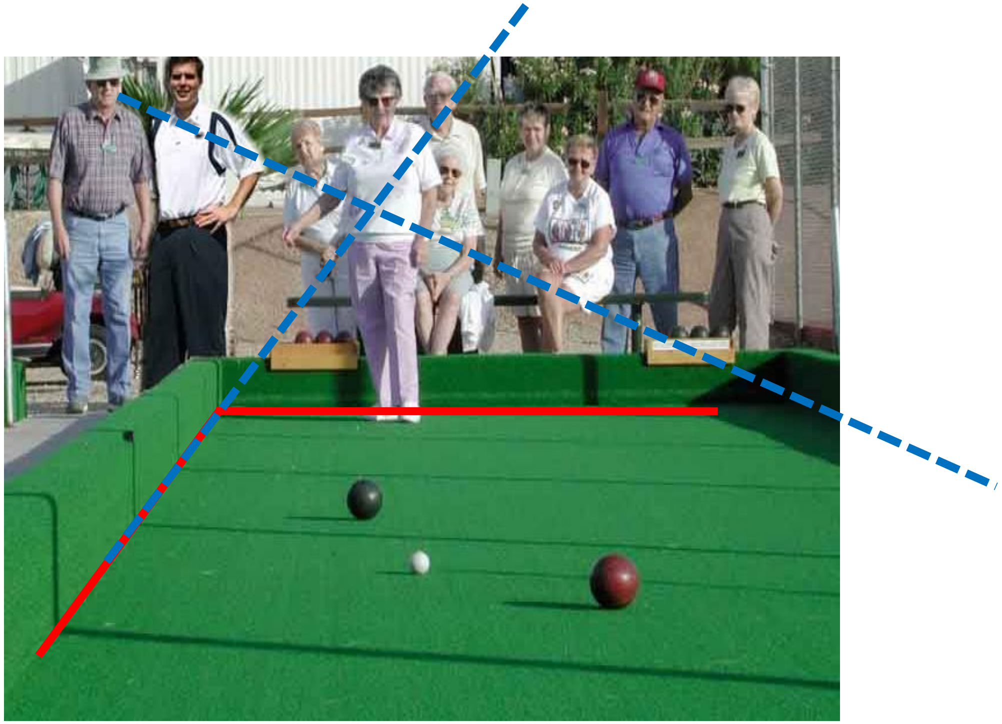
- Length
- Angles



Projective Geometry

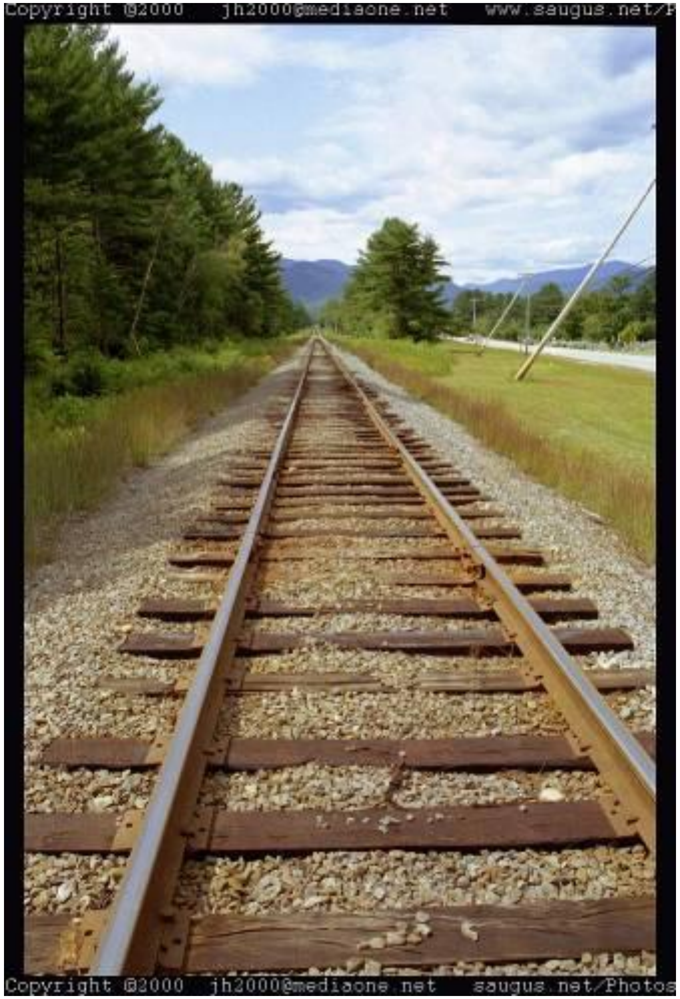
What is preserved?

- Straight lines are still straight



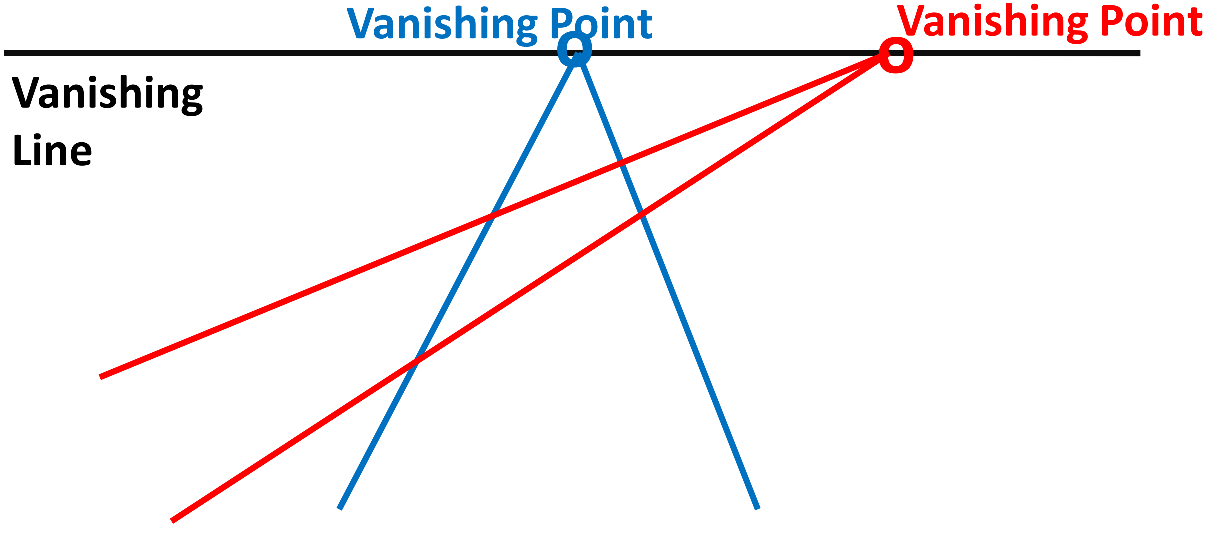
Vanishing points and lines

- Parallel lines in the world intersect in the image at a “vanishing point”

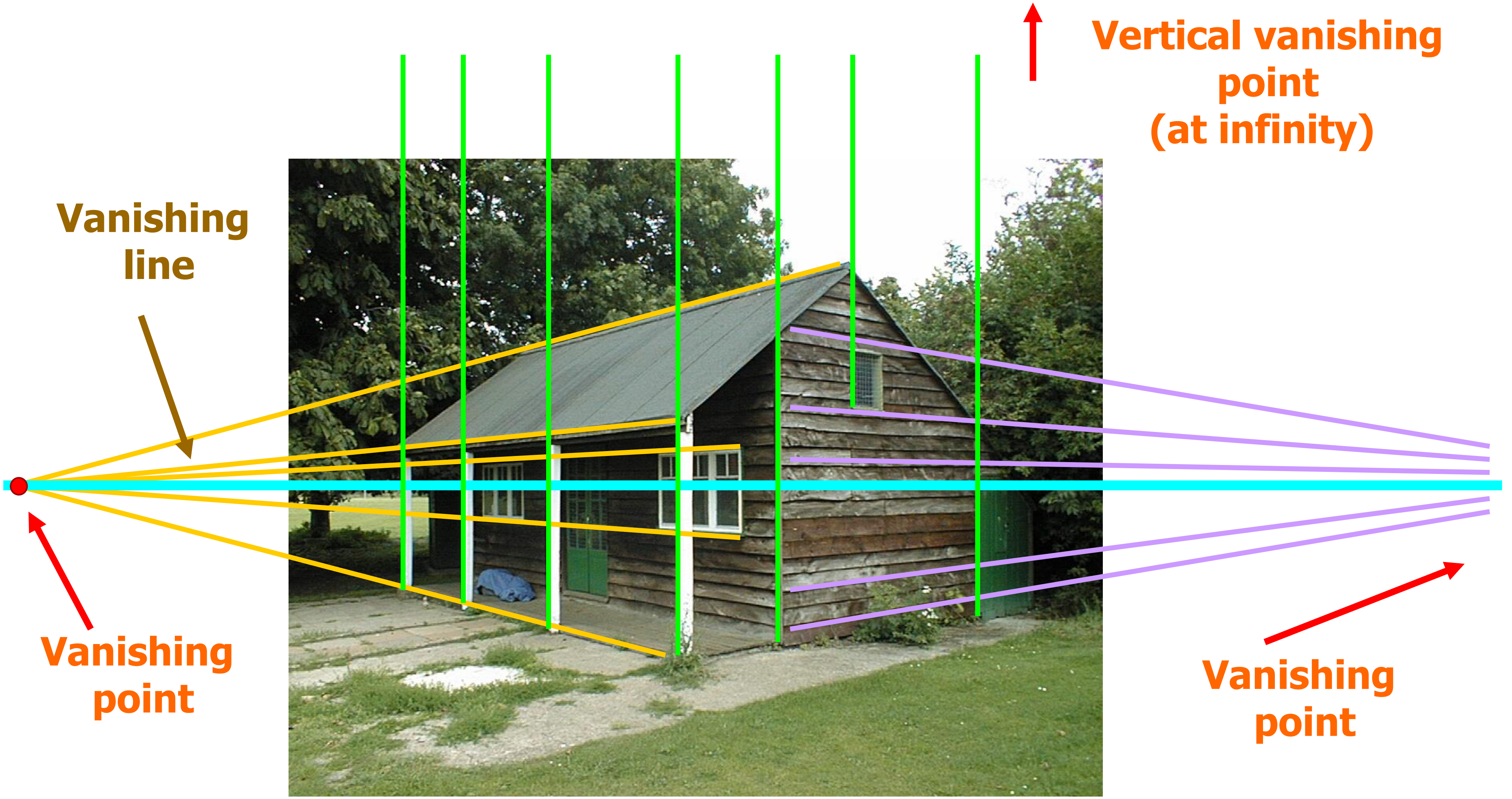


Vanishing points and lines

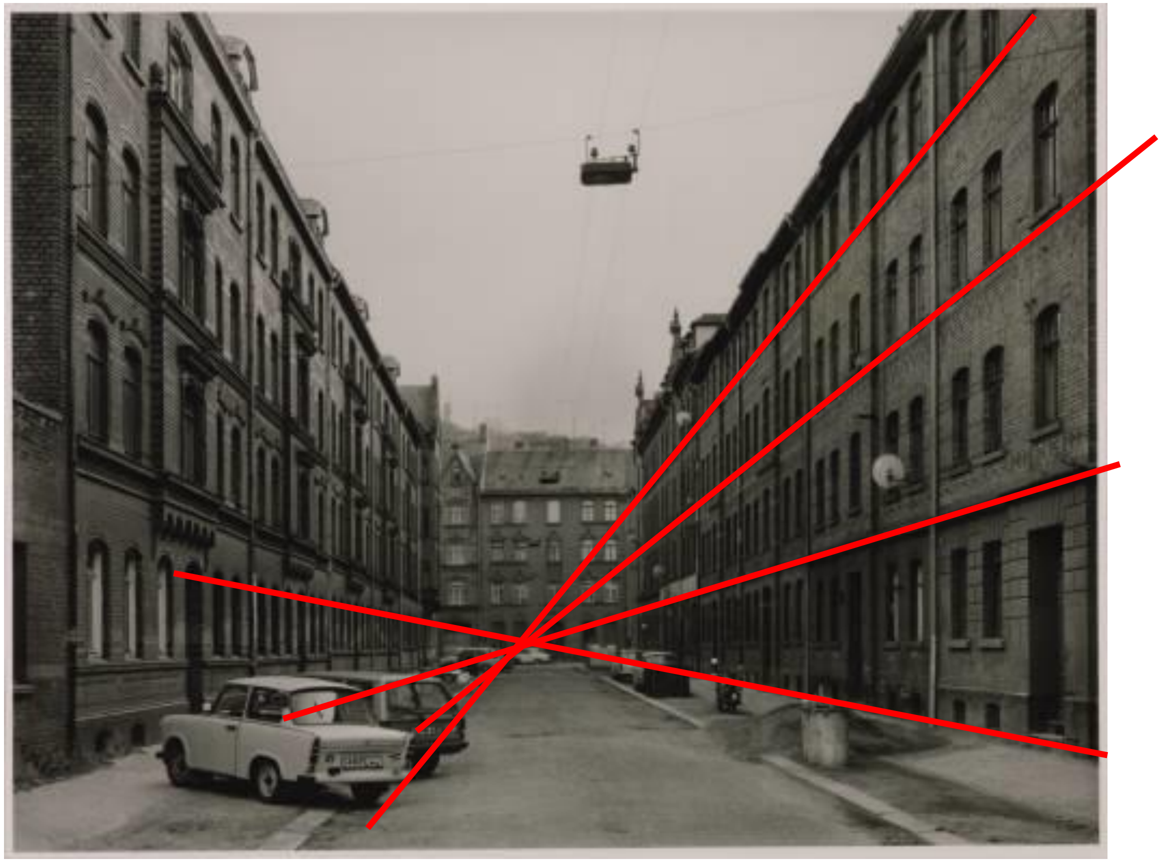
- Parallel lines intersect at a point
- Sets of parallel lines on the same plane form a vanishing line
- Not all lines that intersect are parallel



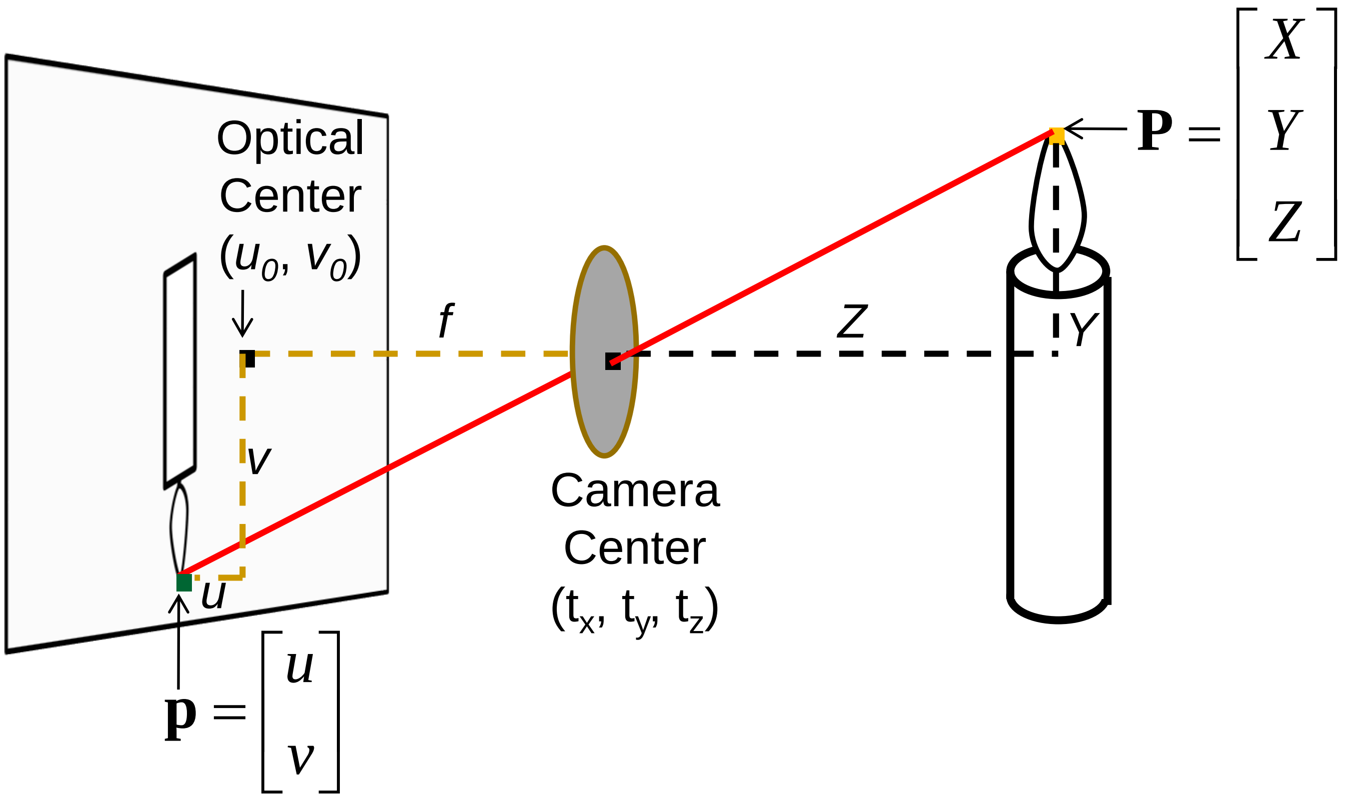
Vanishing points and lines



Vanishing points and lines



Projection: world coordinates→image coordinates



- Recovering information lost by perspective projection is only one, mainly geometric, problem of computer vision

- The second problem is how to understand image brightness.
- The only information available in an intensity image is brightness of the appropriate pixel, which is dependent on a number of independent factors such as

- object surface reflectance properties (given by the surface material, microstructure and marking),
- illumination properties,
- and object surface orientation with respect to a viewer and light source.

- Some scientific and technical disciplines work with 2D images directly; for example,
  - an image of the flat specimen viewed by a microscope with transparent illumination,
  - a character drawn on a sheet of paper,
  - the image of a fingerprint, etc.

- Many basic and useful methods used in digital image analysis do not depend on whether the object was originally 2D or 3D

- Image processing often deals with static images, in which time t is constant.
- A monochromatic （單色）static image is represented by a continuous image function f(x,y) whose arguments are two co-ordinates in the plane

- Computerized image processing uses digital image functions which are usually represented by matrices, so co-ordinates are integer numbers.
- The range of image function values is also limited; by convention, in monochromatic images the lowest value corresponds to black and the highest to white.
- Brightness values bounded by these limits are ***gray levels***.

Image digitization

- ***Sampling*** means measuring the value of an image at a finite number of points.
- ***Quantization*** is the representation of the measured value at the sampled point by an integer.

- An image captured by a sensor is expressed as a continuous function f(x,y) of two co-ordinates in the plane.
- Image digitization means that the function f(x,y) is ***sampled*** into a matrix with M rows and N columns.

- The image ***quantization*** assigns to each continuous sample an integer value.
- The continuous range of the image function $f(x,y)$ is split into K intervals.

- The finer the sampling (i.e., the larger M and N) and quantitation (the larger K) the better the approximation of the continuous image function f(x,y).

Sampling

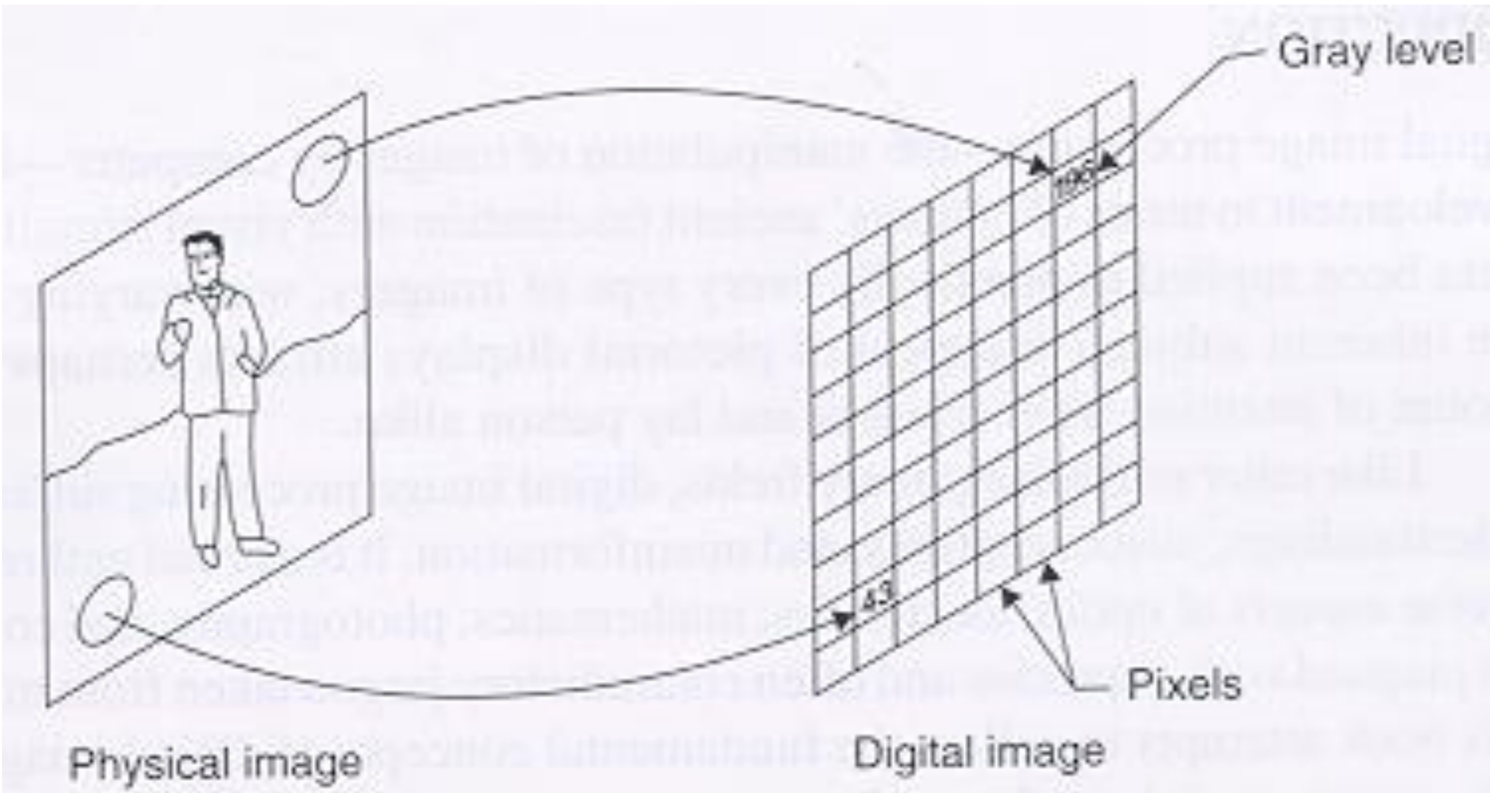


Image digitization (cont'd)

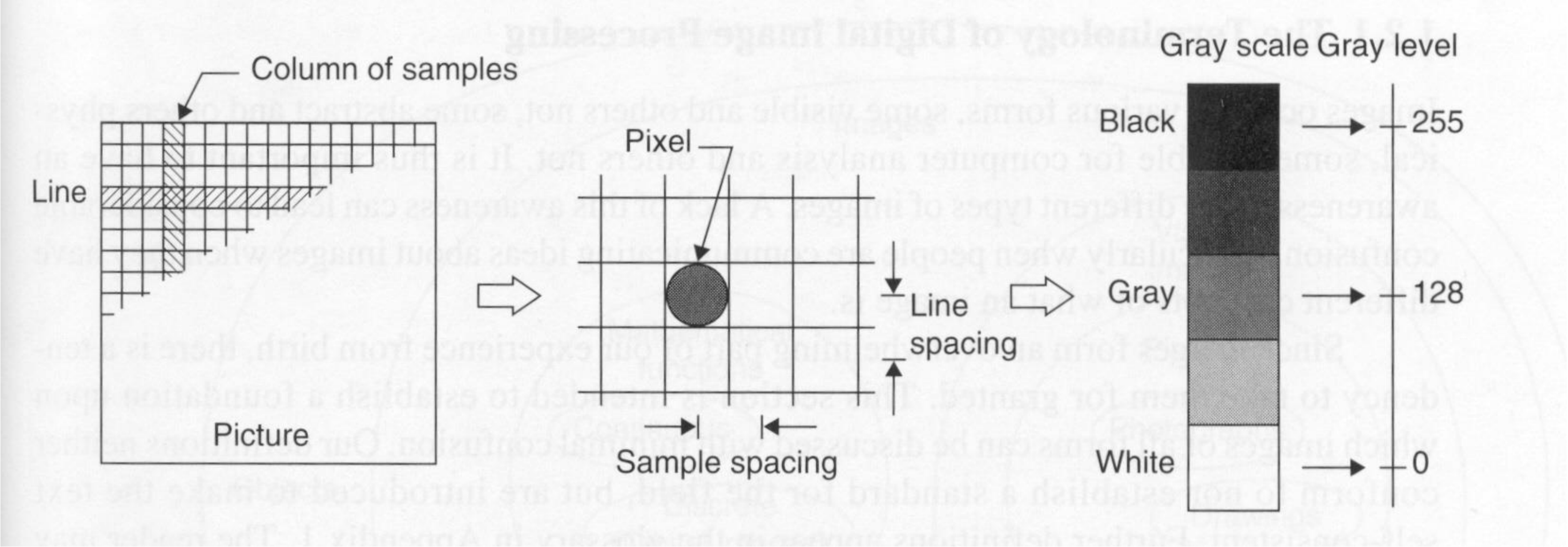


Image quantization (example)

256 gray levels (8bits/pixel)

32 gray levels (5bits/pixel)

16 gray levels (4bits/pixel)

8 gray levels (3bits/pixel)

4 gray levels (2bits/pixel)

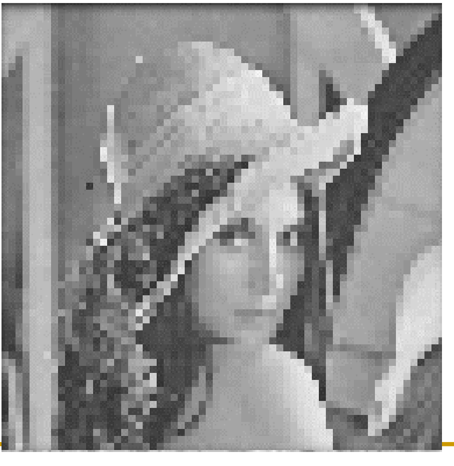
2 gray levels (1bits/pixel)

Image sampling (example)

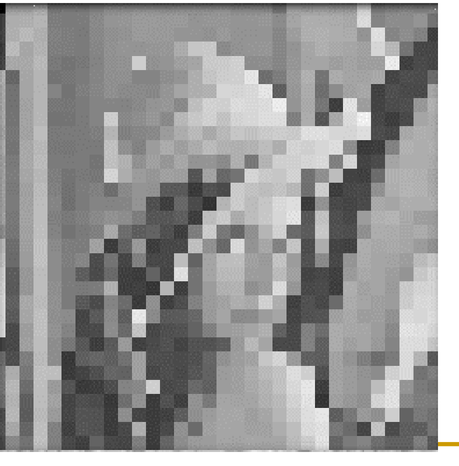
original image

sampled by a factor of 2

sampled by a factor of 4



sampled by a factor of 8



- N: # of rows, M: # of columns, Q: # of gray levels
  - N = 2^n, M = 2^m, Q = 2^q (q is the # of bits/pixel)
  - Storage requirements: N*M*Q (e.g., N=M=1024, q=8, 1MB)

$$
\begin{array}{cccc}
f(0,0) & f(0,1) & \ldots & f(0,M-1) \\
f(1,0) & f(1,1) & \ldots & f(1,M-1) \\
\ldots & \ldots & \ldots & \ldots \\
f(N-1,0) & f(N-1,1) & \ldots & f(N-1,M-1)
\end{array}
$$

Image coordinate system

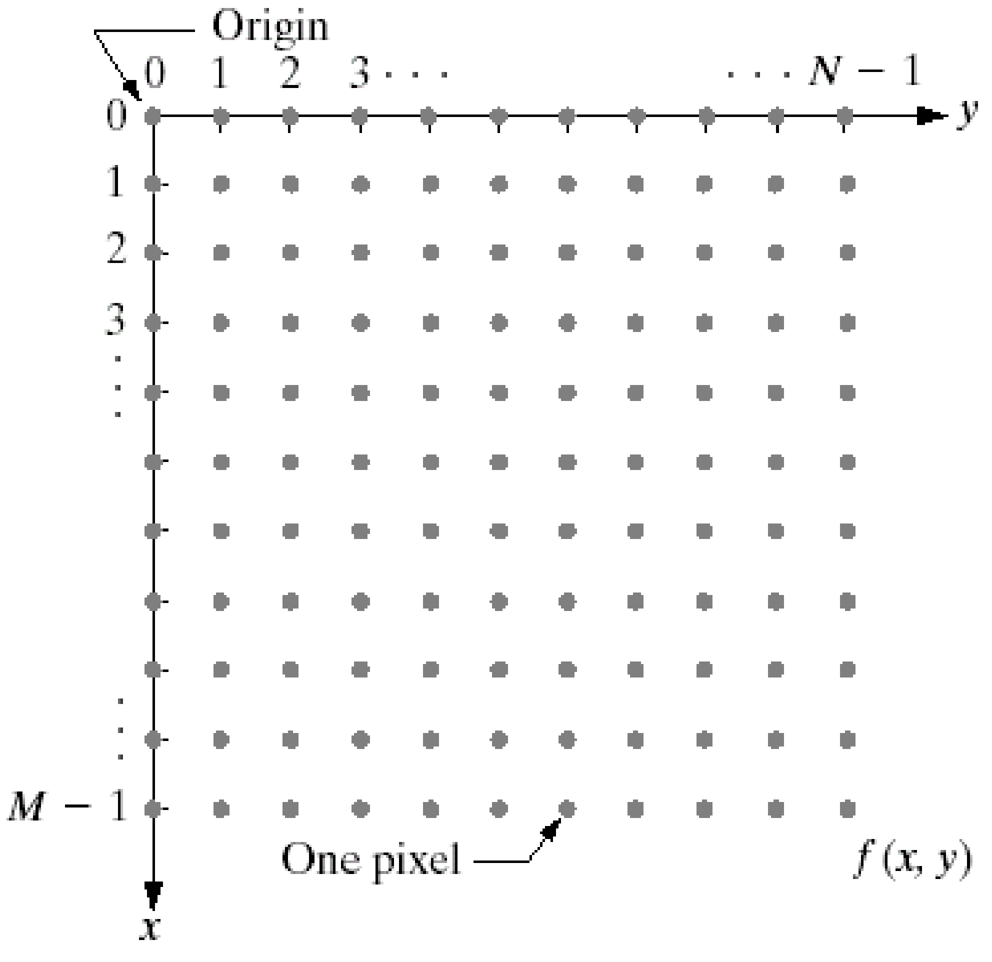
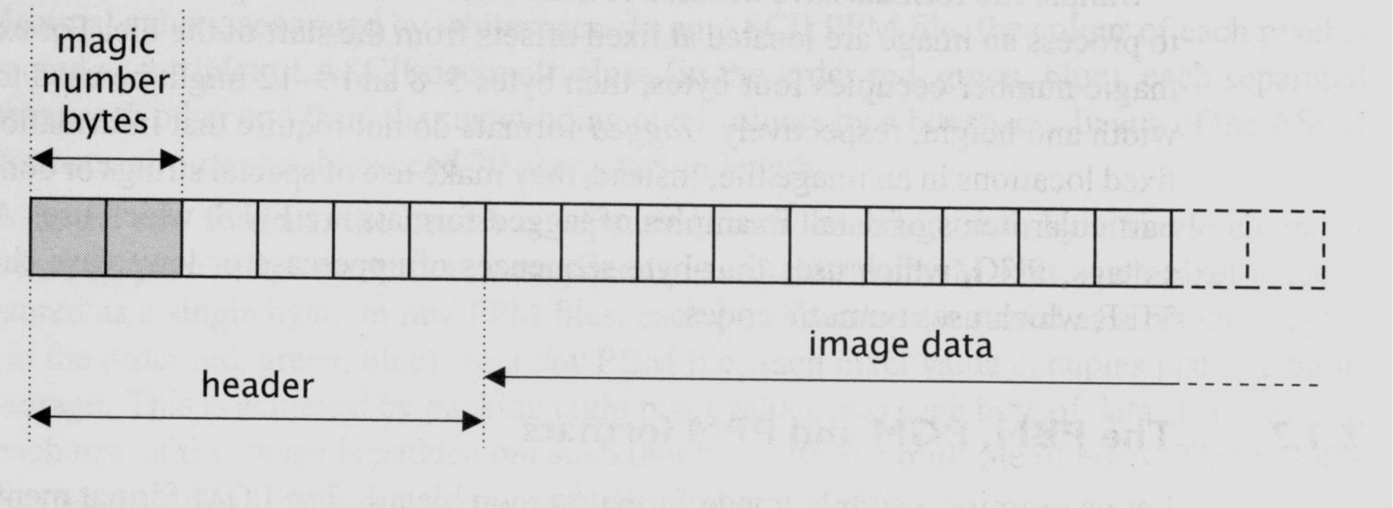


Image file formats

- Many image formats adhere to the simple model shown below (line by line, no breaks between lines).
- The header contains at least the width and height of the image.
- Most headers begin with a ***signature*** or “magic number” - a short sequence of bytes for identifying the file format.


magic
number
bytes
image data
header

Common image file formats

- **GIF** (Graphic Interchange Format): has been used to encode a huge number of images in the internet and databases.
  - Relatively easy to work with
  - Cannot be used for high-precision（高精度）color, since only 8-bits are used to encode color
    - Normally 256 color values available
    - 16-color is another option
  - Non-lossy compression is available

- **TIFF** (Tagged Image File Format): it is used on all popular platforms and is often the format used by scanners.
  - Supports multiple images with 1 to 24 bits of color per pixel.
  - Options are available for either lossy or lossless compression

- **JPEG** (Joint Photographic Experts Group): is a more recent standard from the ***Joint Photographic Experts Group***, the major purpose was to provide for practical compression of high-quality color still images.
  - Allows for real-time hardware for coding and decoding
  - High compression: a flexible but complex lossy coding scheme is used which often can compression a high quality image 20:1 without noticeable degradation.
  - The compression works well when the image has large regions of nearly constant color.
  - Is not designed for video.

- **MPEG** (Motion Picture Experts Group): a stream-oriented encoding scheme for video, audio, text, and graphics.
  - MPEG-1: suitable for multimedia for popular personal computers, but are too low for high-quality TV.
  - MPEG-2: handle high definition TV rates.