Digital Image Processing
Chapter 2
Digital Image Fundamentals

Fall 2018

吳俊霖
Jiunn-Lin Wu
jlwu@cs.nchu.edu.tw

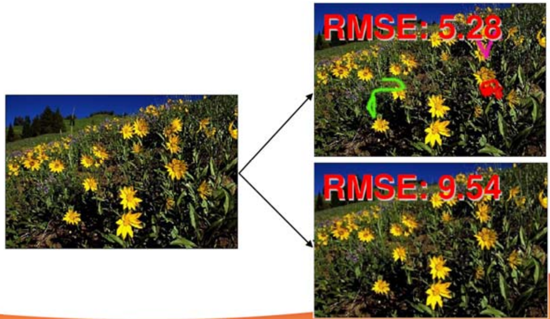
國立中興大學

Elements of Visual Perception

- The importance of visual perception in DIP: Although the foundation of digital image processing is based on mathematical and probabilistic formulations, human subjective visual judgments plays a central role in the choice of one technique over another. 概率公式
  - Tone mapping
  - Image compression (JPEG)
- We are interested in the mechanics and parameters related to how images are formed in our eyes.

Elements of Visual Perception

- Human subjective visual judgments plays a central role in the choice of one technique over another.



Elements of Visual Perception

- Structure of the Human Eye

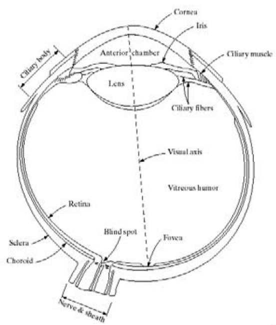


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

Elements of Visual Perception

- Two classes of light receptors
  - Cones
    - They are highly sensitive to color.
    - can resolve fine details because has its own nerve.
  - Rods
    - Lower resolution because several rods share a nerve.
    - The number of rods is much larger.
    - Rods serve to give a general, overall pictures of the fields of view
    - They are not involved in color vision and are sensitive to low levels of illuminations.

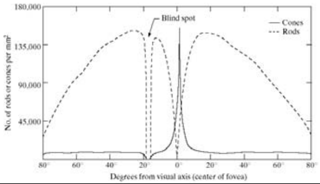
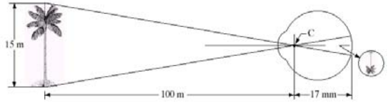


FIGURE 2.2 Distribution of rods and cones in the retina.

Elements of Visual Perception

- Image formation in the eye

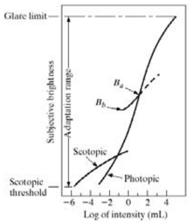
FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.



Elements of Visual Perception

- Brightness and Adaptation and Discrimination
  - Light intensity
  - Subjective brightness
    - Intensity as perceived by the human visual system.
    - It is a logarithmic function of the light intensity incident on the eye.
  - Brightness adaptation

FIGURE 2.4 Range of subjective brightness sensations showing a particular adaptation level.



Elements of Visual Perception

- Brightness and Adaptation and Discrimination
  - The ability of the eye to discriminate between changes in light intensity at any specific adaptation level is considerable interest.
  - Weber ratio $\frac{\Delta I_c}{I}$

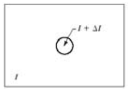
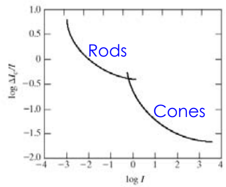


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

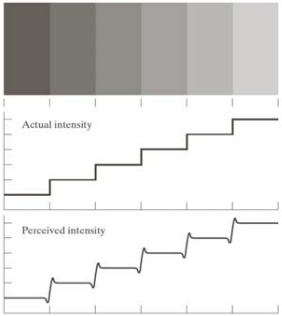
  - The brightness discrimination is poor at low level of illumination.

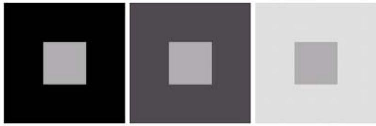
FIGURE 2.6 Typical Weber ratio as a function of intensity.



Elements of Visual Perception

- Brightness and Adaptation and Discrimination
  - The perceived brightness is not a simple function of intensity. (log)

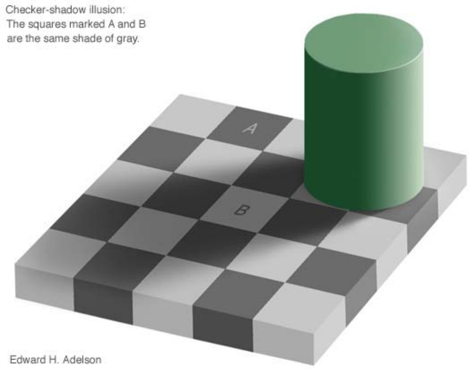




a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Checker-Shadow Illusion

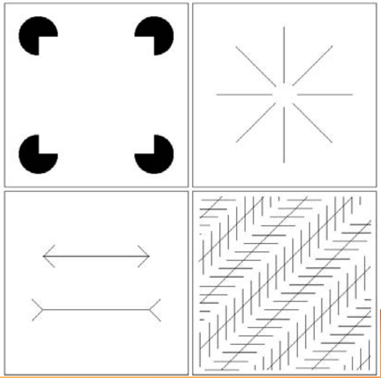


Elements of Visual Perception

- Optical illusions

a b
c d

FIGURE 2.9 Some well-known optical illusions.



Light and the EM Spectrum

- Radio waves have photons with low energies
- Gamma rays have photons with high energies.

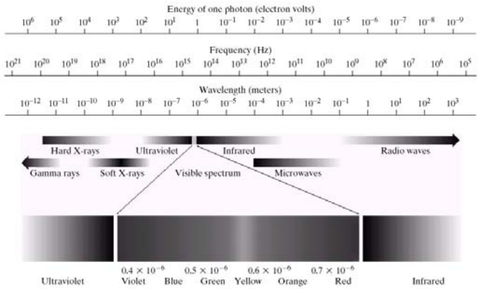


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Light and the EM Spectrum

- Gray level
  - Light that is void of color is called achromatic or monochromatic light.
  - The only attribute of such light is its intensity.
  - The term gray level generally is used to describe monochromatic intensity.
- Radiance
  - It is the total amount of energy that flows from the light source.
  - It is usually measured in watts (W).
- Luminance
  - Gives a measure of the amount of energy an observer perceives from a light source.
  - Measured in lumens (lm)
- Brightness
  - It is a subjective descriptor of light perception that is practically impossible to measure.

Image Sensing and Acquisition

- The wavelength of EM wave required to see an object must be the same size or smaller than the object.
- The images are generated by the combination of an "Illumination" source and the reflection or absorption of energy from that source by the element of the "Scene" being images.
- The energy reflected from or passing through the scene.

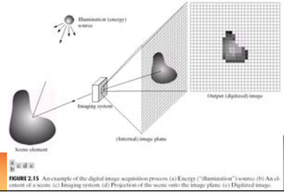


Image Sensing and Acquisition

- Sensors: convert detected energy to electrical voltage signal.

a
b
c
FIGURE 2.12 (a) Single imaging sensor. (b) Line sensor. (c) Array sensor.

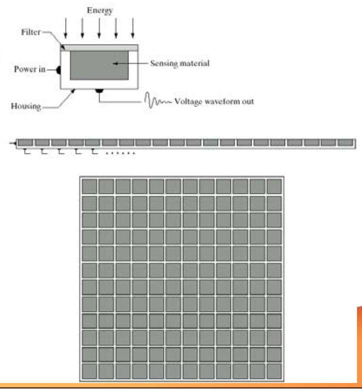


Image Sensing and Acquisition

- Image acquisition using a single sensor
  - This method is an inexpensive but slow way to obtain high-resolution images.

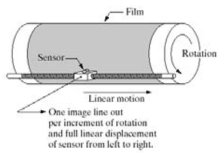
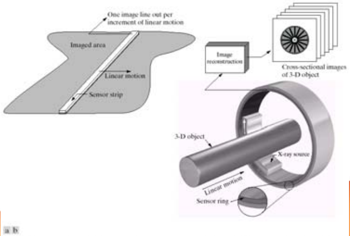


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Image Sensing and Acquisition

- Image acquisition using sensor strips
  - This is the type of arrangement used in most flat bed scanners.
  - Computerized Axial Tomography (CAT): Sensor strips mounted in a ring configuration are used in medial and industrial imaging to obtain cross-sectional ("slice") images of 3-D objects.
    - The output must be processed by reconstruction algorithm.
    - Other modalities of imaging based on CAT: MRI and PET



a b
FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

CCD vs. CMOS

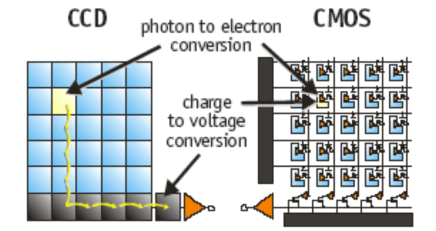


Figure 3. CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node; CMOS imagers convert charge to voltage inside each pixel.

The data is from http://www.dalsa.com/

CCD vs. CMOS

- Feature Comparison

Feature	CCD	CMOS
Signal out of pixel	Electron packet	Voltage
Signal out of chip	Voltage (analog)	Bits (digital)
Signal out of camera	Bits (digital)	Bits (digital)
Fill factor	High	Moderate
Amplifier mismatch	N/A	Moderate
System Noise	Low	Moderate
System Complexity	High	Low
Sensor Complexity	Low	High
Camera components	Sensor + multiple support chips + lens	Sensor + lens possible, but additional support chips common
Relative R&D cost	Lower	Higher
Relative system cost	Depends on Application	Depends on Application

The data is from http://www.dalsa.com/

CCD vs. CMOS

- Performance Comparison

Performance	CCD	CMOS
Responsivity	Moderate	Slightly better
Dynamic Range	High	Moderate
Uniformity	High	Low to Moderate
Uniform Shuttering	Fast, common	Poor
Uniformity	High	Low to Moderate
Speed	Moderate to High	Higher
Windowing	Limited	Extensive
Antiblooming	High to none	High
Biasing and Clocking	Multiple, higher voltage	Single, low-voltage

The data is from http://www.dalsa.com/

CCD vs. CMOS

- CMOS Development's Winding Path

Initial Prediction for CMOS	Twist	Outcome
Equivalence to CCD in imaging performance	Required much greater process adaptation and deeper submicron lithography than initially thought	High performance available in CMOS, but with higher development cost than CCD
On-chip circuit integration	Longer development cycles, increased cost, tradeoffs with noise, flexibility during operation	Greater integration in CMOS, but companion chips still required for both CMOS and CCD
Reduced power consumption	Steady improvement in CCDs	Advantage for CMOS, but margin diminished
Reduced imaging subsystem size	Optics, companion chips and packaging are often the dominant factors in imaging subsystem size	CCDs and CMOS comparable
Economies of scale from using mainstream logic and memory foundries	Extensive process development and optimization required	CMOS imagers use legacy production lines with highly adapted processes akin to CCD fabrication

The data is from http://www.dalsa.com/

CMOS Sensors

- Canon has used CMOS sensors in most of its digital cameras for years, including the professional line of digital SLRs.
- Nikon D2X CMOS sensor made by Sony
  - According to Japanese website Nikkei Electronics Online the 12.4 megapixel CMOS sensor used in the Nikon D2X is made by Sony. (2004/9/17)
- Nikon has today lifted the covers on their first full-frame digital SLR, the new 12.1 megapixel D3. The D3 is all about speed and sensitivity, twelve megapixels on a big CMOS chip means large photosites (8.45 µm pitch to be precise) and that adds up to base sensitivity of ISO 200 to 6400. (2007/8/23)
- In addition to the FX format (Full-frame) D3, Nikon has also announced the new DX format D300 which also features a 12 megapixel CMOS sensor, 14-bit A/D conversion, EXPEED image processor, sensitivity up to ISO 6400 (with boost).
- Sony is introducing its full-frame α (alpha) DSLR-A900 camera, aimed at serious photo enthusiasts looking for traditional SLR performance with the added benefits of digital photography. It is designed to deliver ultra-fine picture quality with the world's highest resolution, 24.6-megapixel, 35 mm full-frame CMOS sensor and fast image processing with a new dual BIONZ® processing engines. The camera is also the first to have a body-integrated image stabilization system. (Sept. 9, 2008)

Image Sensing and Acquisition

- Image acquisition using sensor arrays
  - This is the predominant arrangement found in digital cameras.

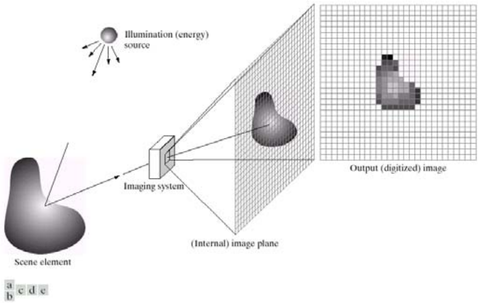


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Image Sensing and Acquisition

- A Simple Image Formation Model
  - f(x,y) must be nonzero and finite. $0 < f(x,y) < \infty$
  - f(x,y) is characterized by two components.
    - Illumination
    - Reflection

$$f(x,y) = i(x,y)r(x,y) \qquad 0 < i(x,y) < \infty \qquad 0 < r(x,y) < 1$$

  - Gray scale
    - Shift the interval [*Lmin*, *Lmax*] → [0, L-1]
    - 0 is black; L-1 is white

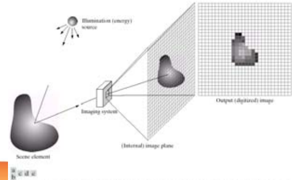
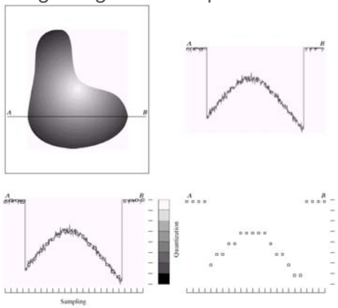


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Sampling and Quantization

- Sampling: digitizing the spatial coordinates values.
- Quantization: digitizing of the amplitude values.

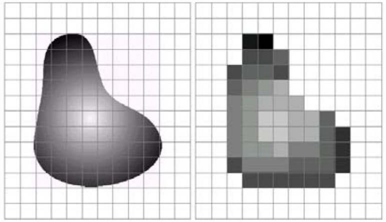


a b
c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Sampling and Quantization

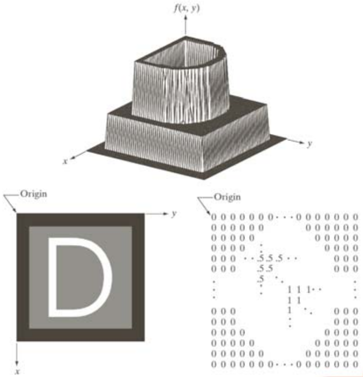
- The image quality is determined by the number of samples and discrete gray levels used in sampling and quantizations.



a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Representing Digital Images



a
b c

FIGURE 2.18 (a) Image plotted as a surface. (b) Image displayed as a visual intensity array. (c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

Sampling and Quantization

- Representing Digital Images
  - Pixel

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

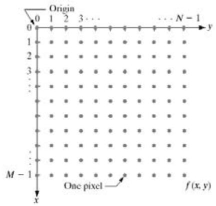


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

$$\mathbf{A}=\begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}$$

Saturation and Noise

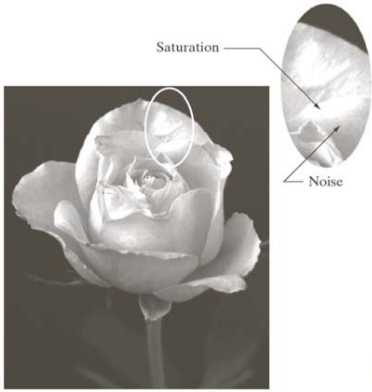


FIGURE 2.19 An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity levels are clipped (note how the entire saturated area has a high, *constant* intensity level). Noise in this case appears as a grainy texture pattern. Noise, especially in the darker regions of an image (e.g., the stem of the rose) masks the lowest detectable true intensity level.

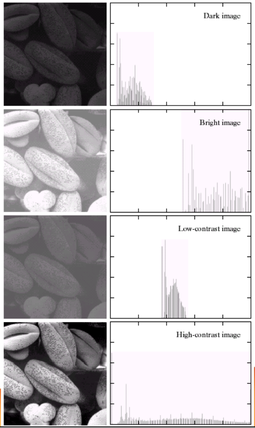
Sampling and Quantization

- *N,M,L* are positive integer and *L* is power of 2.
- [0,L-1] is called "dynamic range".
  - Images whose gray levels span a significant portion of the gray scale as having a high dynamic range.
  - When an appreciable number of pixels exhibit the above property, the image will have high contrast.
- Storage of a digitized image is $b = M \times N \times k$ where $L = 2^k$

TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

High Contrast Image



What is Dynamic Range?

- For a nature image: the ratio of the lightest to darkest point in the image.
- For a camera: the ratio of the intensities that just saturate the camera and the adjust lift the camera response cone standard deviation above the camera noise.
- For a display: the ratio of the maximum and minimum image intensities emitted from the screen.

The Dynamic Range Problem

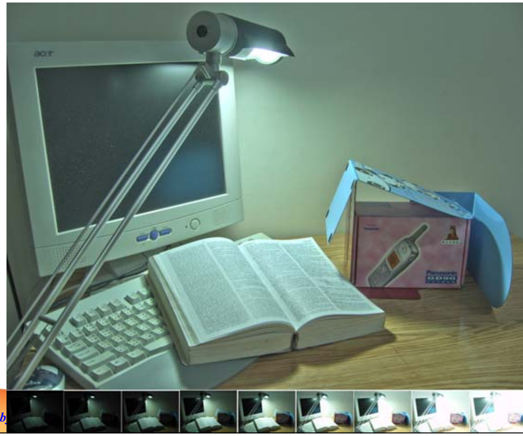
- Scene Photo


The Dynamic Range Problem

- Scene Photo


HDR Image Rendering:

HDR Image Rendering:



Spatial Resolution

- Intuitively, spatial resolution is a measure of the smallest discernible detail in an image.
- Quantitatively, spatial resolution can be stated in a number of ways, with *line pairs per unit distance*, and dots (pixels) per distance being among the most common measures.
- Dots per unit distance is a measure of image resolution used commonly in the printing and publishing industry.
- In the U.S., this measure usually is expressed as ***dots per inch*** **(dpi)**.

Spatial Resolution

- Sampling is the principle factor determining the spatial resolution of an image.

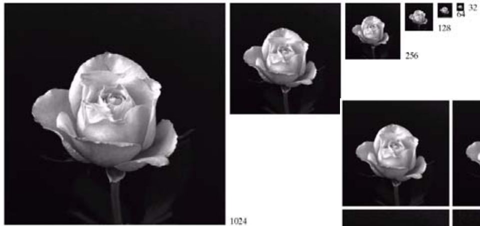
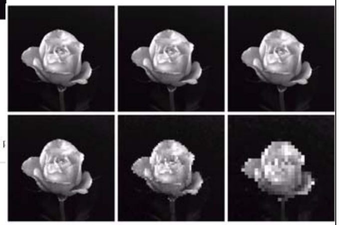


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 p gray levels was kept at 256.



a b c
d e f

FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

Spatial Resolution

a b
c d

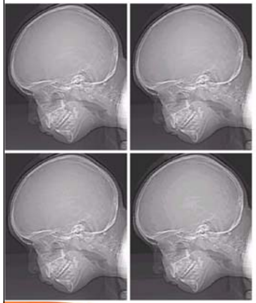
FIGURE 2.20 Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.

Intensity Resolution

- Intensity resolution (gray level resolution) similarly refers to the smallest discernible change in intensity level.
- It is common practice to refer to the number of bits used to quantize intensity as the *intensity resolution*.
- Based on hardware consideration, the number of intensity levels usually is an integer power of two.
- Sometimes one finds systems that can digitize the intensity levels of an image using 10-12 bits, but these are the exception, rather than the rule.
- Canon redefines the future of photography: 21.1 MP EOS 5D Mark II offers Full HD video capture. (17 September 2008)
  - **New CMOS sensor:**
    The EOS 5D Mark II's newly designed full frame 21.1 Megapixel CMOS sensor features ISO sensitivity from 100-6400, expandable to 50, 12,800 and 25,600.
  - **New DIGIC 4 processor**
    A new DIGIC 4 processor combined with 14-bit analogue to digital conversion provides smooth gradations in mono-tonal areas such as skies, and highly accurate colour rendition.

Intensity Resolution

- Intensity resolution (Gray level resolution)
  - False contour

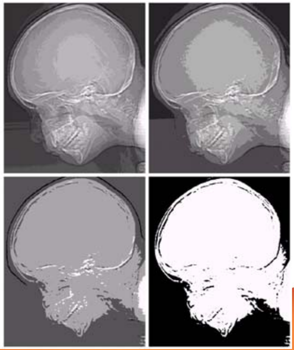


a b
c d

FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

e f
g h

FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Spatial and Intensity Resolution

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

Spatial and Intensity Resolution

- It suggests that for images with a large amount of detail only a few gray levels may be needed.

FIGURE 2.23 Representative isopreference curves for the three types of images in Fig. 2.22.

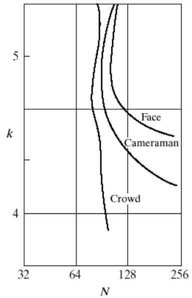


Image Interpolation

- Nearest neighbor interpolation
- Bilinear interpolation
- Bicubic interpolation

a b c
d e f

FIGURE 2.24 (a) Image reduced to 72 dpi and zoomed back to its original size (3692 × 2812 pixels) using nearest neighbor interpolation. This figure is the same as Fig. 2.20(d). (b) Image shrunk and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation. (d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. 2.24(d) is the same as Fig. 2.20(c)]. Compare Figs. 2.24(e) and (f), especially the latter, with the original image in Fig. 2.20(a).

Sampling and Quantization

- Alising and Moire pattern

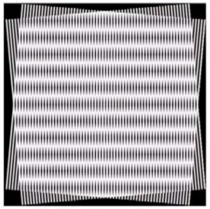


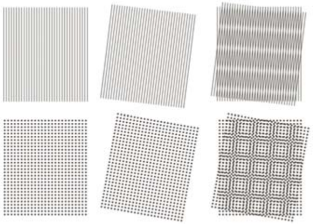
FIGURE 2.24 Illustration of the Moiré pattern effect.

Aliasing

a b c

FIGURE 4.17 Illustration of aliasing on resampled images. (a) A digital image with negligible visual aliasing. (b) Result of resizing the image to 50% of its original size by pixel deletion. Aliasing is clearly visible. (c) Result of blurring the image in (a) with a 3 × 3 averaging filter prior to resizing. The image is slightly more blurred than (b), but aliasing is not longer objectionable. (Original image courtesy of the Signal Compression Laboratory, University of California, Santa Barbara.)

Moiré Pattern



a b c
d e f

FIGURE 4.20 Examples of the moiré effect. These are ink drawings, not digitized patterns. Superimposing one pattern on the other is equivalent mathematically to multiplying the patterns.

The Effect of Undersampling: Aliasing

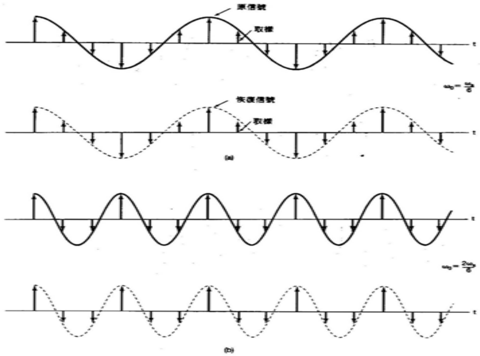


圖 7.16 頻疊效應對正弦信號的影響。對不同的 ω_0，原先正弦信號（實線）的取樣及其還原信號（虛線）如圖所示：(a) $\omega_0 = \omega_s/6$；(b) $\omega_0 = 2\omega_s/6$；(c) $\omega_0 = 4\omega_s/6$；(d) $\omega_0 = 5\omega_s/6$。在(a)和(b)沒有頻疊產生，然而在(c)和(d)有頻疊產生

The Effect of Undersampling: Aliasing

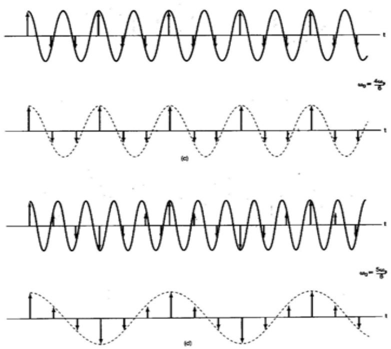
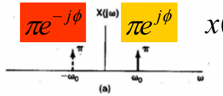


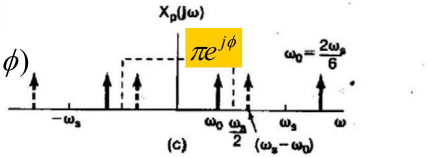
圖 7.16 (續)

Phase Reversal

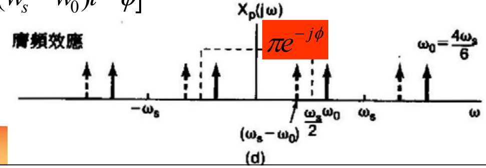


$$x(t) = \cos(w_0 t + \phi)$$

$$x_r(t) = \cos(w_0 t + \phi)$$



$$x_r(t) = \cos[(w_s - w_0)t - \phi]$$



Moiré Pattern

- If a scene contains areas with repetitive detail which exceeds the resolution of the camera
  - a wavy moiré pattern can appear, as shown in crop A.
  - There is no moiré in crop B of an image of the same scene taken with a camera with a higher resolution.
  - Anti-alias filters reduce or eliminate moiré but also reduce image sharpness.
- The data is from Vincent Bockaert's website.

A

B

Bayer CFA Pattern

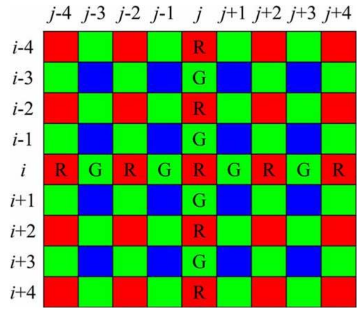


Figure: A 9x9 window of Bayer CFA (Color Filter Array) pattern.

Demosaicing

- King-Hong Chung and Yuk-Hee Chan, *"Color Demosaicing Using Variance of Color Differences"*, *IEEE TRANS. ON IMAGE PROCESSING,* VOL. 15, NO. 10, OCTOBER 2006, pp. 2944-2955.
  - This paper presents an adaptive demosaicing algorithm.
  - Missing green samples are first estimated based on the variances of the color differences along different edge directions. The missing red and blue components are then estimated basedon the interpolated green plane.
  - This algorithm can effectively preserve the details in texture regions and, at the same time, it can significantly reduce the color artifacts.
  - As compared with the latest demosaicing algorithms, the proposed algorithm produces the best average demosaicing performance both objectively and subjectively.

Demosaicing Results

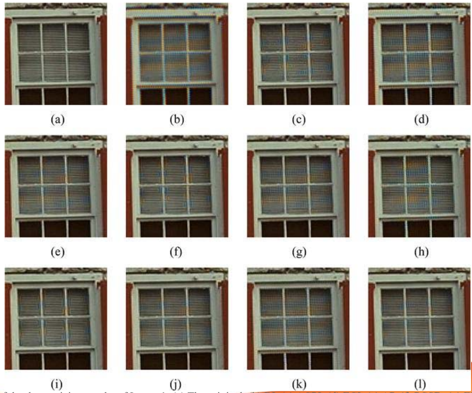


Fig. 6. Part of the demosaicing results of Image 1: (a) The original, (b) BI, (c) ACPI, (d) ECI, (e) AP, (f) PCSD, (g) EECI, (h) DUOR, (i) AHDDA, (j) DSA, (k) DAFD, and (l) the proposed algorithm.

Demosaicing Results

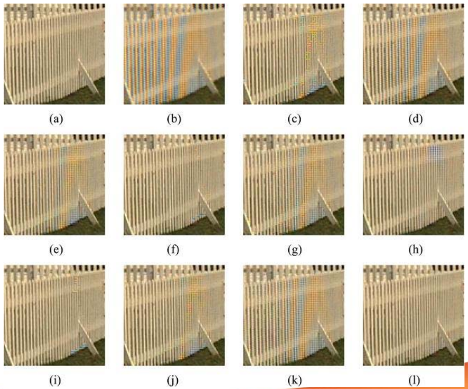
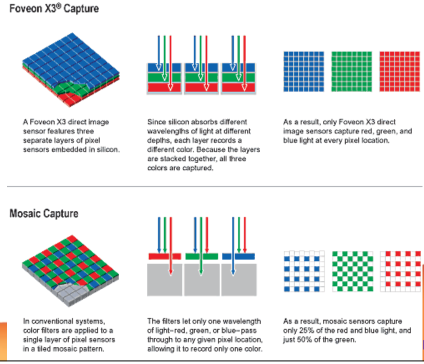


Fig. 8. Part of the demosaicing results of Image 19: (a) The original, (b) BI, (c) ACPI, (d) ECI, (e) AP, (f) PCSD, (g) EECI, (h) DUOR, (i) AHDDA, (j) DSA, and (k) the proposed algorithm.

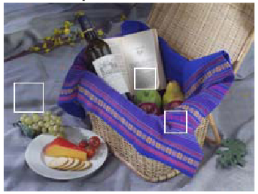
Foveon X3 Technology

- The Foveon X3 direct image sensor is the most advanced color image sensor ever developed. (Sigma DSLR use X3).
- Why X3 is better?



Foveon X3 Technology-Image Comparison

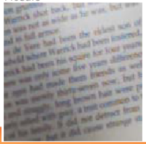
Mosaic Capture



Foveon X3 Capture

Sharpness

Mosaic



Foveon X3

Foveon X3 Technology-Image Comparison

Color Detail

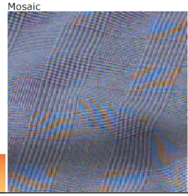
Mosaic

Foveon X3

[Larger View]

Artifacts

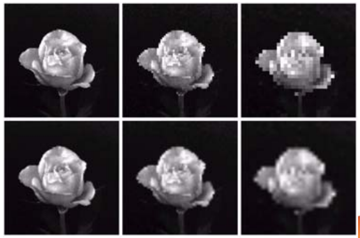
Mosaic



Foveon X3

Sampling and Quantization

- Zooming and Shrinking Digital Images
  - Nearest neighbor interpolation (pixel replication)
    - may cause checkerboard effects
  - Bilinear interpolation
  - Bicubic interpolation



a b c
d e f

FIGURE 2.25 Top row: images zoomed from 128×128, 64×64, and 32×32 pixels to 1024×1024 pixels, using nearest neighbor gray-level interpolation. Bottom row: same sequence, but using bilinear interpolation.

Fractal Image Compression

- Higher compression ratio
- The detail can be created at all scales.

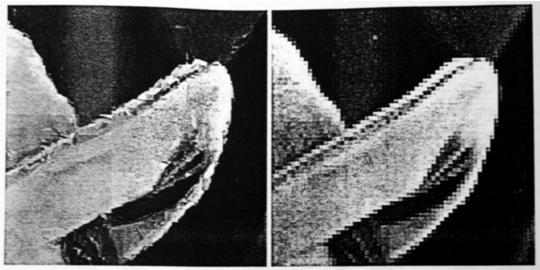
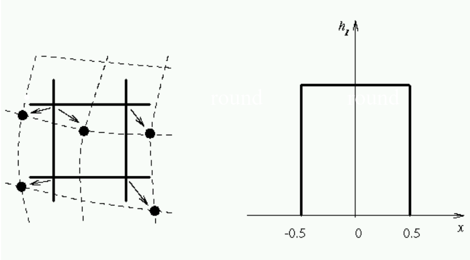


Figure 1.4: A portion of Lenna's hat decoded at 4 times its encoding size (left), and the original image enlarged to 4 times its size (right), showing pixelization.

Nearest neighbor interpolation

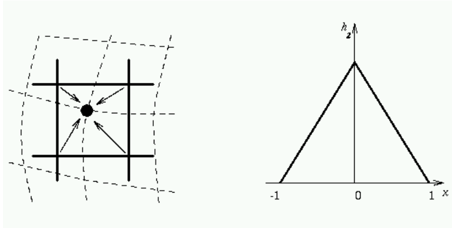
- Assign to the point the brightness value of the nearest point g in the discrete raster.



$$f_1(x,y) = g_s(round(x), round(y))$$

Bilinear interpolation

- Bilinear interpolation explores four points neighboring the point (x,y), and assumes that the brightness function is bilinear in this neighborhood.



$$f_2(x,y) = (1-a)(1-b)g_s(l,k) + a(1-b)g_s(l+1,k) + (1-a)bg_s(l,k+1) + abg_s(l+1,k+1)$$

Bi-cubic interpolation

- Bi-cubic interpolation improves the model of the brightness function by approximating it locally by a bicubic polynomial surface; 16 neighboring points are used for interpolation.
- Interpolation kernel (`Mexican hat') is defined via

$$h(t) = \begin{cases} 1-2|t|^2+|t|^3, & if\ |t|<1 \\ 4-8|t|+5|t|^2-|t|^3, & if\ 1 \le |t| < 2 \\ 0, & otherwise \end{cases}$$

Sampling and Quantization

- Zooming and Shrinking Digital Images
  - To reduce possible aliasing effects, it is a good idea to blur an image slightly before shrinking it.
  - Gaussian Pyramid

Fig. 2a. *The Gaussian pyramid. The original image, G_0, is repeatedly filtered and subsampled to generate the sequence of reduced resolution image G_1, G_2, etc. These comprise a set of lowpass-filtered copies of the original image in which the bandwidth decreases in one-octave steps.*

Sampling and Quantization

- Gaussian Pyramid

Fig. 2b. Levels of the Gaussian pyramid expanded to the size of the original image. The effects of lowpass filtering are now clearly apparent.

Some Basic Relationships between Pixels

- Neighbors of a pixel
  - 4-neighbors

    $$N_4 = (x+1, y), (x-1, y), (x, y+1), (x, y-1)$$

  - Diagonal neighbors

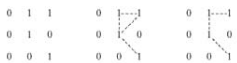
    $$N_D = (x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

  - 8-neighbors

    $$N_8 = N_4 + N_D$$

Some Basic Relationships between Pixels

- Adjacency, Connectivity, Regions and Boundaries
  - 4-adjacency: two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$
  - 8-adjacency: two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$
  - m-adjacency: two pixels p and q with values from V are m-adjacent if
    - q is in $N_4(p)$,or
    - q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.



a b c

FIGURE 2.26 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

Some Basic Relationships between Pixels

- Path
  - 4-, 8-, or m-paths
  - Length
  - Closed path
- Connectivity
  - Connected
    - Two pixels p and q are said to be connected in S if there existed a path between them consisting entirely of pixels in S.
  - Connected component
    - For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S.
  - Connected set
    - If it only has one connected component then set S is called a connected set.

Some Basic Relationships between Pixels

- Regions
  - We call R a region of the image if R is a connected set.
- Boundary (border/contour)
  - The boundary of a region R is the set of pixels in the region that have one or more neighbors that are not in R.
  - The borer of a region is the set of pixels in the region that have at least one background neighbor.
  - The preceding definition sometimes is referred to as the *inner border* of the region to distinguish it from its *outer border*, which is the corresponding border in the background .

Some Basic Relationships between Pixels

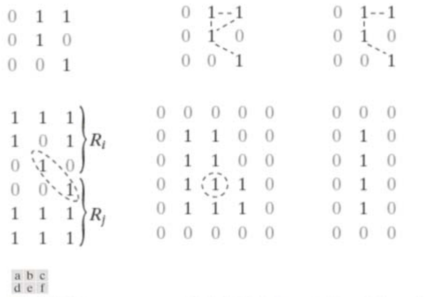


FIGURE 2.25 (a) An arrangement of pixels. (b) Pixels that are 8-adjacent (adjacency is shown by dashed lines; note the ambiguity). (c) *m*-adjacency. (d) Two regions that are adjacent if 8-adjecency is used. (e) The circled point is part of the boundary of the 1-valued pixels only if 8-adjacency between the region and background is used. (f) The inner boundary of the 1-valued region does not form a closed path, but its outer boundary does.

Some Basic Relationships between Pixels

- Distance measures
  - Distance function or metric
    - $D(p,q) \geq 0 \quad (D(p,q) = 0 \text{ iff } p = q),$
    - $D(p,q) = D(q,p), and$
    - $D(p,z) \leq D(p,q) + D(q,z).$
- Enclidean distance: $D_e(p,q) = \left[(x-s)^2 + (y-t)^2\right]^{\frac{1}{2}}$
- D_4 distance (city-block distance): $D_4(p,q) = |x-s| + |y-t|$
- D_8 distance (chessboard distance): $D_8(p,q) = \max(|x-s|,|y-t|)$

An Introduction to the Mathematical Tools Used in DIP

- Array versus Matrix Operations
- Linear versus Nonlinear Operations
- Arithmetic Operations
- Set and Logic Operations
- Spatial Operations
- Vector and Matrix Operations
- Image Transforms
- Probabilistic Methods

Linear versus Nonlinear Operations

- Linear operations

$$x_1[n] \rightarrow y_1[n]$$

$$x_2[n] \rightarrow y_2[n]$$

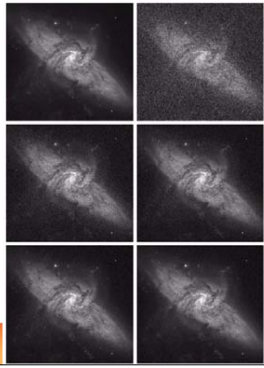
$$ax_1[n] + bx_2[n] \rightarrow ay_1[n] + by_2[n]$$

- Nonlinear operations

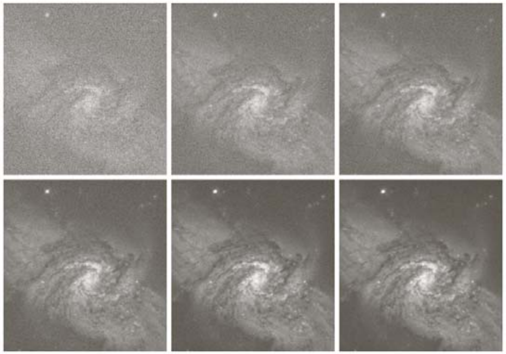
Enhancement Using Arithmetic Operations

- Image averaging
  - a noisy image $g(x,y) = f(x,y) + \eta(x,y)$
  - $\bar{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$
  - $E\{\bar{g}(x,y)\} = f(x,y)$

$$\sigma^2_{\bar{g}(x,y)} = \frac{1}{K}\sigma^2_{\eta(x,y)}$$



Arithmetic Operations

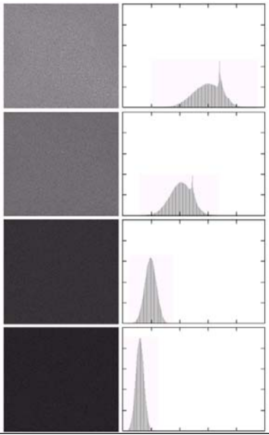


a b c
d e f

FIGURE 2.26 (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)–(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)

Enhancement Using Arithmetic/Logic Operations

- Image averaging



a b

FIGURE 3.31 (a) From top to bottom: Difference images between Fig. 3.30(a) and the four images in Figs. 3.30(c) through (f), respectively. (b) Corresponding histograms.

SONY-手持夜景拍攝模式

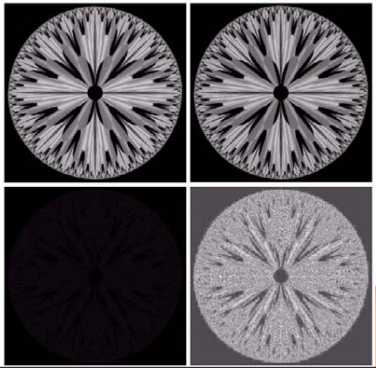
- 即使調整高 ISO 設定或未使用腳架拍攝，手持夜景拍攝模式也能呈現更清晰、銳利的影像。
- 在手持夜景拍攝模式下，最高能降低 50% 的影像雜訊。這是因為高感光度的「Exmor R」CMOS 感光元件與 BIONZ 影像處理器的結合，使 Cyber-shot WX1 能重疊六張栩栩如生的畫面並建立單張最佳影像。

Enhancement Using Arithmetic Operations

- A frequent application of image subtraction is in the enhancement of differences between images.

a b
c d

FIGURE 3.28 (a) Original fractal image. (b) Result of setting the four lower-order bit planes to zero. (c) Difference between (a) and (b). (d) Histogram-equalized difference image. (Original image courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA).



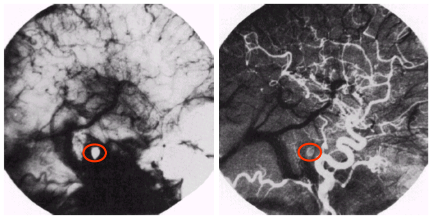
Enhancement Using Arithmetic Operations

a b c

FIGURE 2.27 (a) Infrared image of the Washington, D.C. area. (b) Image obtained by setting to zero the least significant bit of every pixel in (a). (c) Difference of the two images, scaled to the range [0, 255] for clarity.

Enhancement Using Arithmetic/Logic Operations

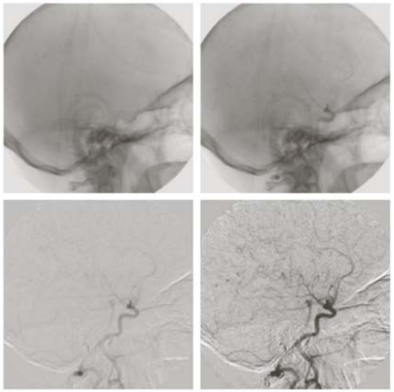
- Image subtraction
  - Mask mode radiography



a b

FIGURE 3.29 Enhancement by image subtraction. (a) Mask image. (b) An image (taken after injection of a contrast medium into the bloodstream) with mask subtracted out.

Arithmetic Operations



a b
c d

FIGURE 2.28 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

Image Sequence Analysis

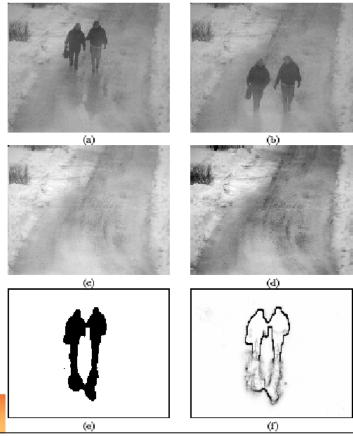


Image Sequence Analysis

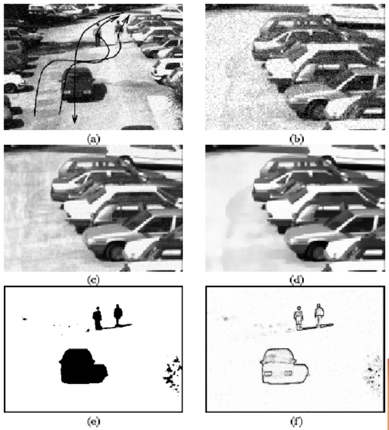
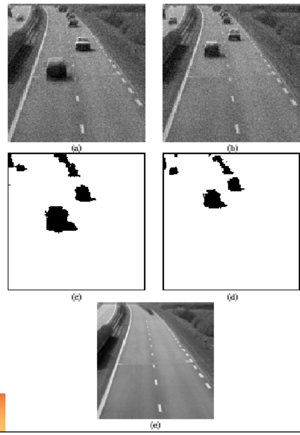


Image Sequence Analysis



Arithmetic Operations

- An important application of image multiplication (and division) is *shading correction*.

a b c

FIGURE 2.29 Shading correction. (a) Shaded SEM image of a tungsten filament and support, magnified approximately 130 times. (b) The shading pattern. (c) Product of (a) by the reciprocal of (b). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

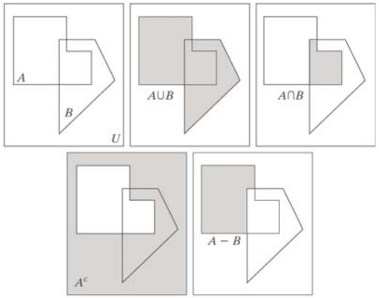
Arithmetic Operations

- Another common use of image multiplication is in *masking*, also called *region of interest (ROI)*, operations.

a b c

FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

Set Operations

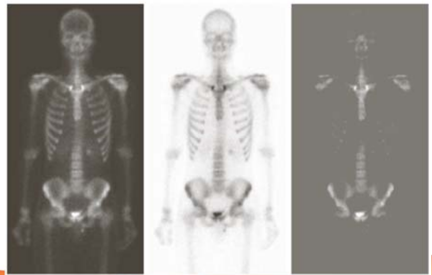


a b c
d e

FIGURE 2.31 (a) Two sets of coordinates, A and B, in 2-D space. (b) The union of A and B. (c) The intersection of A and B. (d) The complement of A. (e) The difference between A and B. In (b)–(e) the shaded areas represent the member of the set operation indicated.

Set Operations

- $A_n = A^c = \{(x, y, 255 - z) \mid (x, y, z) \in A\}$
- $A \cup B = \left\{ \max_z(a, b) \mid a \in A, b \in B \right\}$



a b c

FIGURE 2.32 Set operations involving gray-scale images. (a) Original image. (b) Image negative obtained using set complementation. (c) The union of (a) and a constant image. (Original image courtesy of G.E. Medical Systems.)

Set Operations

- Composition of the firework photos

Logical Operations on Binary Images

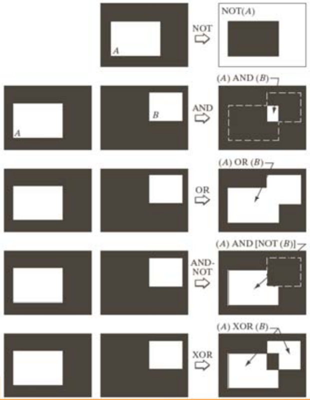
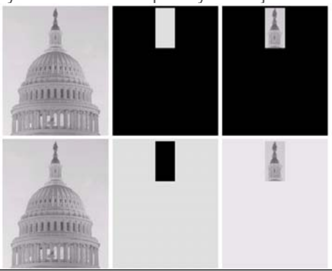


FIGURE 2.33 Illustration of logical operations involving foreground (white) pixels. Black represents binary 0s and white binary 1s. The dashed lines are shown for reference only. They are not part of the result.

Enhancement Using Logic Operations

- Arithmetic/logic operations involving images are performed on a pixel-by pixel basis between two or more images.
- AND, OR, an NOT are functionally complete
  - The AND an OR operations are used for masking (ROI).
  - They will be used frequently in conjunction with morphology.



a b c
d e f

FIGURE 3.27 (a) Original image. (b) AND image mask. (c) Result of the AND operation on images (a) and (b). (d) Original image. (e) OR image mask. (f) Result of operation OR on images (d) and (e).

Spatial Operations

- Spatial operations are performed directly on the pixels of a given image.
  - Single-pixel operations
  - Neighborhood operations
  - Geometric spatial transformations
- Single-pixel operations $s = T(z)$

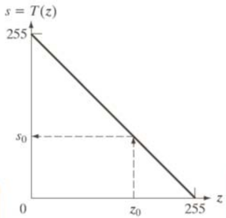
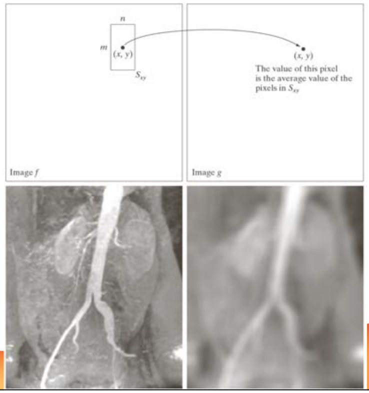


FIGURE 2.34 Intensity transformation function used to obtain the negative of an 8-bit image. The dashed arrows show transformation of an arbitrary input intensity value z_0 into its corresponding output value s_0.

Spatial Operations

- Neighborhood operations



a b
c d

FIGURE 2.35 Local averaging using neighborhood processing. The procedure is illustrated in (a) and (b) for a rectangular neighborhood. (c) The aortic angiogram discussed in Section 1.3.2. (d) The result of using Eq. (2.6-21) with $m = n = 41$. The images are of size 790×686 pixels.

Geometric Spatial Transforms

- Geometric transformations modify the spatial relationship between pixels in an image.
- Geometric transformations often are called *rubber-sheet transformations*.
- A geometric transformation consists of two basic operations:
  - *Spatial transformation*, which defines the "re-arrangement" of pixels on the image plane;
  - *Gray-level interpolation*, which deals with the assignment of gray levels in the spatially transformed image.

Geometric Spatial Transforms

- One of the most commonly used spatial coordinate transformations is the affine transform, which has the general form:

$$[x \ \ y \ \ 1] = [v \ \ w \ \ 1]\,\mathbf{T} = [v \ \ w \ \ 1]\begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix}$$

- In practice, we can use the above equation in two basic ways;
  - Forward mapping
  - Inverse mapping

Affine Transformations

TABLE 2.2
Affine transformations based on Eq. (2.6.–23).

Transformation Name	Affine Matrix, T	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = w$	
Scaling	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = c_x v$ $y = c_y w$	
Rotation	$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v\cos\theta - w\sin\theta$ $y = v\cos\theta + w\sin\theta$	
Translation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$	$x = v + t_x$ $y = w + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v + s_v w$ $y = w$	
Shear (horizontal)	$\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = s_h v + w$	

Geometric spatial transforms

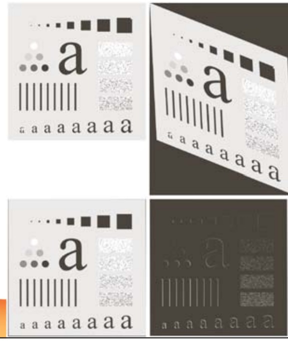
- The objective of this example is to illustrate image rotation using an affine transform.

a b c d

FIGURE 2.36 (a) A 300 dpi image of the letter T. (b) Image rotated 21° clockwise using nearest neighbor interpolation to assign intensity values to the spatially transformed pixels. (c) Image rotated 21° using bilinear interpolation. (d) Image rotated 21° using bicubic interpolation. The enlarged sections show edge detail for the three interpolation approaches.

Image Registration

- Image Registration is an important application of digital image processing used to align two or more images of the same scene.



a b
c d

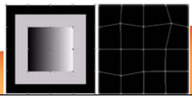
FIGURE 2.37 Image registration. (a) Reference image. (b) Input (geometrically distorted image). Corresponding tie points are shown as small white squares near the corners. (c) Registered image (note the errors in the borders). (d) Difference between (a) and (c), showing more registration errors.

Geometric Spatial Transforms & Image Registration

- An *image f* with pixel coordinates *(x,y)* undergoes geometric distortion to produce an *image g* with coordinates *(x',y')*.

$$x' = r(x, y)$$
$$y' = s(x, y)$$

- Formulating a single set of analytical functions *r(x,y)* and *s(x,y)* that describe the geometric distortion process over the entire image plane generally is not possible.
- The method used most frequently to overcome this difficulty is to formulate the spatial relocation of pixels by the use of *tiepoints*, which are a subset of pixels whose location in the input (distorted) and output (corrected) image is known precisely.



Spatial Transformations

- Tiepoints

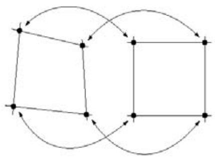


FIGURE 5.32 Corresponding tiepoints in two image segments.

- *Eight parameter method*: Suppose that the geometric distortion process within the quadrilateral regions is modeled by a pair of linear equations so that

$$x' = c_1 x + c_2 y + c_3 xy + c_4$$
$$y' = c_5 x + c_6 y + c_7 xy + c_8$$

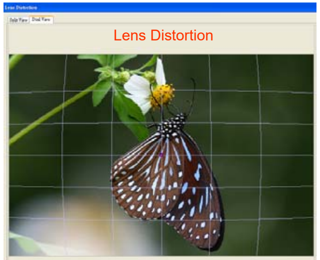
Spatial Transformations

- <u>*Six parameter method*</u>:

$$x' = c_1 x + c_2 y + c_3$$
$$y' = c_4 x + c_5 y + c_6$$

- <u>*Four parameter method*</u>:

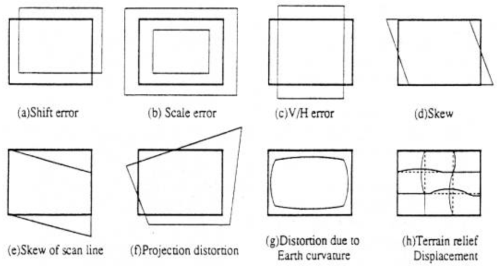
$$x' = c_1 x + c_2 y + c_3$$
$$y' = -c_2 x + c_1 y + c_4$$



- In general, enough tiepoints are needed to generate a set of quadrilateral that over the entire image, with each quadrilateral having its own set of coefficients.

遙測影像幾何變形

- This data is from 中央研究院-計算中心-空間資訊技術小組



遙測影像幾何校正-常用坐標轉換模型

- This data is from 中央研究院-計算中心-空間資訊技術小組

常用坐標轉換模型	適用於大比例尺地圖
Helmert Transform	正形轉換 4 參數
Affine Transform	仿射轉換 6 參數
Projection Transform	投影轉換 8 參數
d. 2' Conformal	二次正形轉換 6 參數

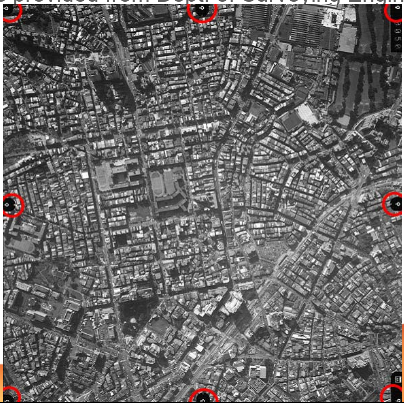
Name	Transform formula	Number of unknown parameters
1) Helmert Transform (scale, shift and rotation)	$x = au + bv + c$ $y = -bu + av + d$	4
2) Affine Transform	$x = au + bv + c$ $y = du + ev + f$	6
3) Pseudo Affine	$x = a_1 uv + a_2 u + a_3 v + a_4$ $y = a_5 uv + a_6 u + a_7 v + a_8$	8
4) Projection Transform	$x = \frac{a_1 u + a_2 v + a_3}{a_7 u + a_8 + 1}$ $y = \frac{a_4 u + a_5 v + a_6}{a_7 u + a_8 + 1}$	8
5) Second-order Conformal	$x = a_1 u + a_2 v + a_3 (u^2 - v^2) + 2a_4 uv + a_5$ $y = -a_2 u + a_1 v + 2a_3 uv - a_4 (u^2 - v^2) + a_6$	6
6) Polynomials	$x = \Sigma\Sigma a_{ij} u^{i-1} v^{j-1}$ $y = \Sigma\Sigma b_{ij} u^{i-1} v^{j-1}$	——

Spatial Transformations

- Tiepoints are established by a number of different techniques, depending on the application. For instance, some image generation system having physical artifacts (such as metallic points) embedded on the imaging sensor itself. These produce a known set of points (called <u>*reseau marks*</u>) directly on the image as it is acquired.
- Remote sensing imaging system
  - Control points (框標)
  - Ground control points (地面控制點)

Remote Sensing Image

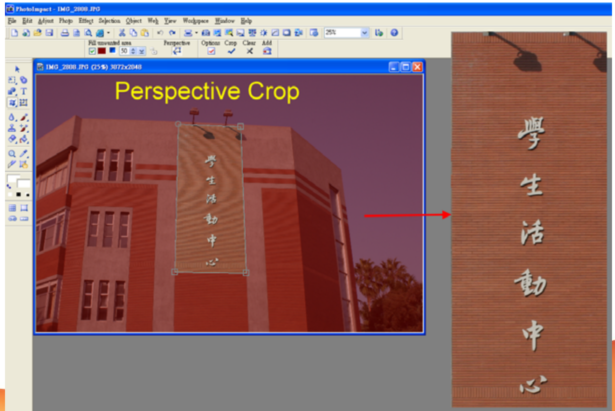
- The image is provided from Dept. of Surveying Engineering, NCKU.



Ground Control Points

- 大坪山, 點位編號:N003 施測時間:84年6月, 標石號碼:無, 標石種類:花崗石, 點位來源:原一等三角點 所在地:台北縣深坑鄉
- 七星山, 點位編號:N001 施測時間:84年6月, 標石號碼:無 標石種類:花崗石, 點位來源:原一等三角點 所在地:台北市北投區
- The data is from http://www.survey.taipei.gov.tw

Perspective Crop



Panorama Factory

- Single-row Stitching – Manual

Gray-Level Interpolation

- The simplest scheme for gray-level interpolation is based on a nearest neighbor approach. This method, also called zero-order interpolation.

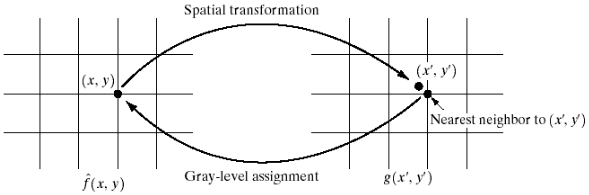
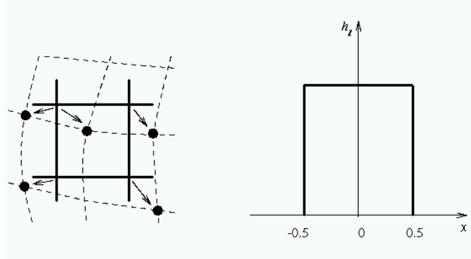


FIGURE 5.33 Gray-level interpolation based on the nearest neighbor concept.

Nearest Neighbor Interpolation

- Assign to the point the brightness value of the nearest point g in the discrete raster.



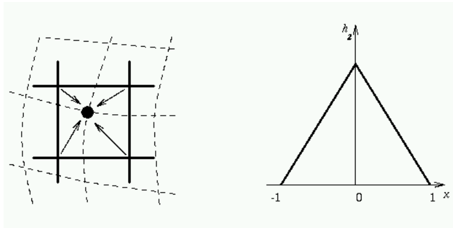
$$f_1(x, y) = g_s(round(x), round(y))$$

Gray-Level Interpolation

- *Cubic convolution interpolation*
  - It fits a surface of the sin(z)/z type through a much larger number of neighbors (say, 16) in order to obtain a smooth estimate of the gray level at any desired point.
  - Typical area in which smoother approximations generally are required include 3-D graphics and medial imaging.
  - The price paid for smoother approximations is additional computational burden.
- For general-purpose image processing a *bilinear interpolation* approach that use the gray levels of the four nearest neighbors usually is adequate.

Bilinear Interpolation

- Bilinear interpolation explores four points neighboring the point (x,y), and assumes that the brightness function is bilinear in this neighborhood.



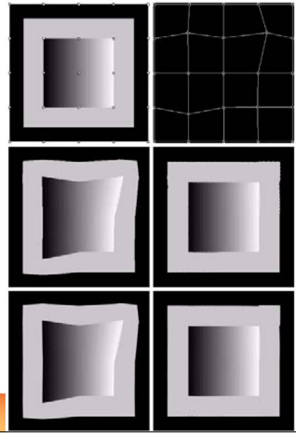
$$f_2(x, y) = (1-a)(1-b)g_s(l, k) + a(1-b)g_s(l+1, k) + (1-a)bg_s(l, k+1) + abg_s(l+1, k+1)$$

Bi-cubic Interpolation

- Bi-cubic interpolation improves the model of the brightness function by approximating it locally by a bicubic polynomial surface; 16 neighboring points are used for interpolation.
- Interpolation kernel (`Mexican hat') is defined via

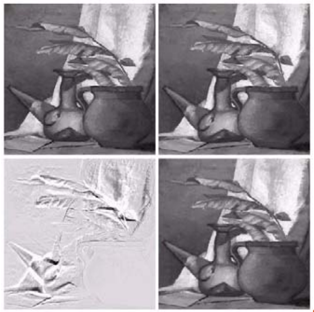
$$h(t)=\begin{cases} 1-2|t|^2+|t|^3, & if\,|t|<1 \\ 4-8|t|+5|t|^2-|t|^3, & if\,1\le |t|<2 \\ 0, & otherwise \end{cases}$$

Geometric Transformations



Geometric Transformations

- When images have more texture, geometric correction errors tend to be less noticeable.



a b
c d

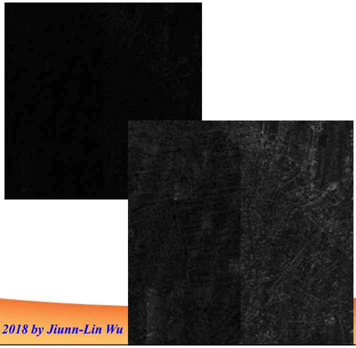
FIGURE 5.35 (a) An image before geometric distortion. (b) Image geometrically distorted using the same parameters as in Fig. 5.34(e). (c) Difference between (a) and (b). (d) Geometrically restored image.

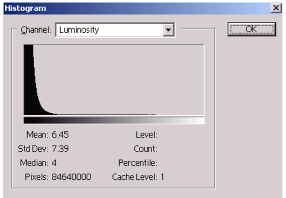
Geometric Transformations

- The image is provided from Dept. of Surveying Engineering, NCKU.

Geometric Transformations

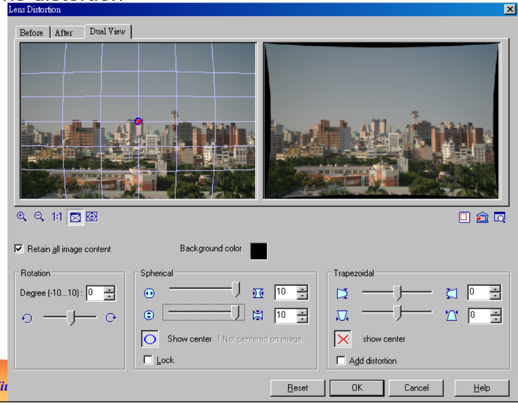
- Six parameter transformation
- The difference image





Geometric Transformations

- Lens distortion



Vector and Matrix Operations

- Multispectral image processing is a typical area in which vector and matrix operations are used routinely.

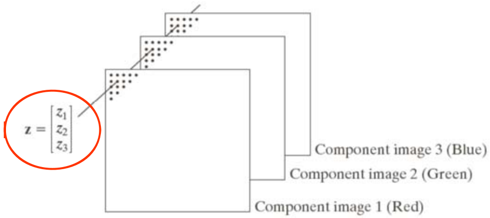


FIGURE 2.38 Formation of a vector from corresponding pixel values in three RGB component images.

Image Transforms

- All the image processing approaches discussed thus far operate directly on the pixel of the input image; that is they work directly in the spatial domain.
- In some cases, image processing tasks are best formulated by transforming the input images, carrying the specified task in a transform domain, and applying the inverse transform to return to the spatial domain.
  - Image compression: JPEG(DCT domain), JPEG2000(Wavelets domain)
  - Image enhancement: Frequency (Fourier) domain

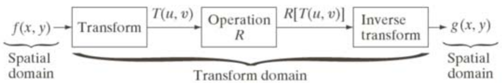
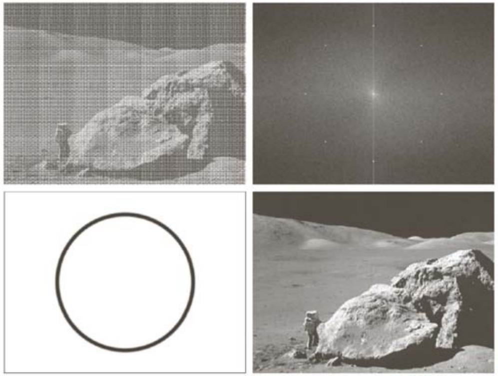


FIGURE 2.39 General approach for operating in the linear transform domain.

Image Transforms



a b
c d

FIGURE 2.40 (a) Image corrupted by sinusoidal interference. (b) Magnitude of the Fourier transform showing the bursts of energy responsible for the interference. (c) Mask used to eliminate the energy bursts. (d) Result of computing the inverse of the modified Fourier transform. (Original image courtesy of NASA.)

Probabilistic Methods

- The probability of intensity level z_k occurring in a given image is estimated as

$$p(z_k) = \frac{n_k}{MN} \qquad \sum_{k=0}^{L-1} p(z_k) = 1$$

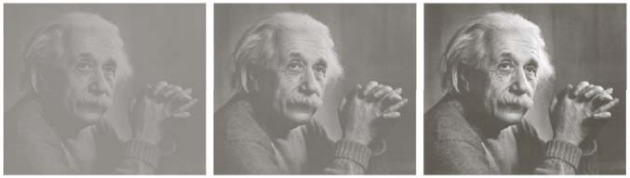
- The mean and variance of the intensities are given by

$$m = \sum_{k=0}^{L-1} z_k p(z_k) \qquad \sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k)$$

- In general, the nth moment of random variable z about the mean is defined as

$$\mu_n(z) = \sum_{k=0}^{L-1} (z_k - m)^n p(z_k)$$

Probabilistic Methods



a b c

FIGURE 2.41 Images exhibiting (a) low contrast, (b) medium contrast, and (c) high contrast.