



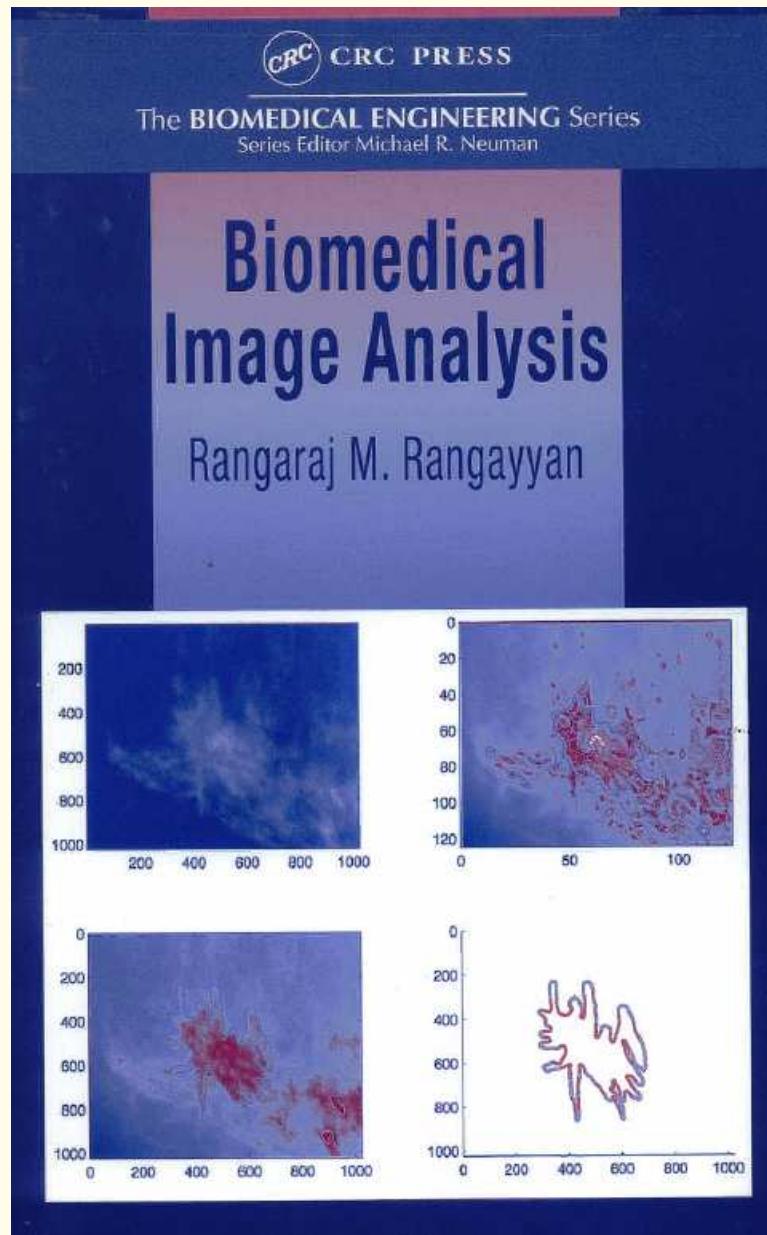
ENEL 697 DIGITAL IMAGE PROCESSING

Rangaraj M. Rangayyan

Professor, Department of Electrical and Computer Engineering
Schulich School of Engineering

Adjunct Professor, Departments of Surgery and Radiology
University of Calgary
Calgary, Alberta, Canada T2N 1N4
Phone: +1 (403) 220-6745
e-mail: ranga@ucalgary.ca

Web: <http://www.enel.ucalgary.ca/People/Ranga/enel697>



CRC Press LLC, Boca Raton, FL, 2005.

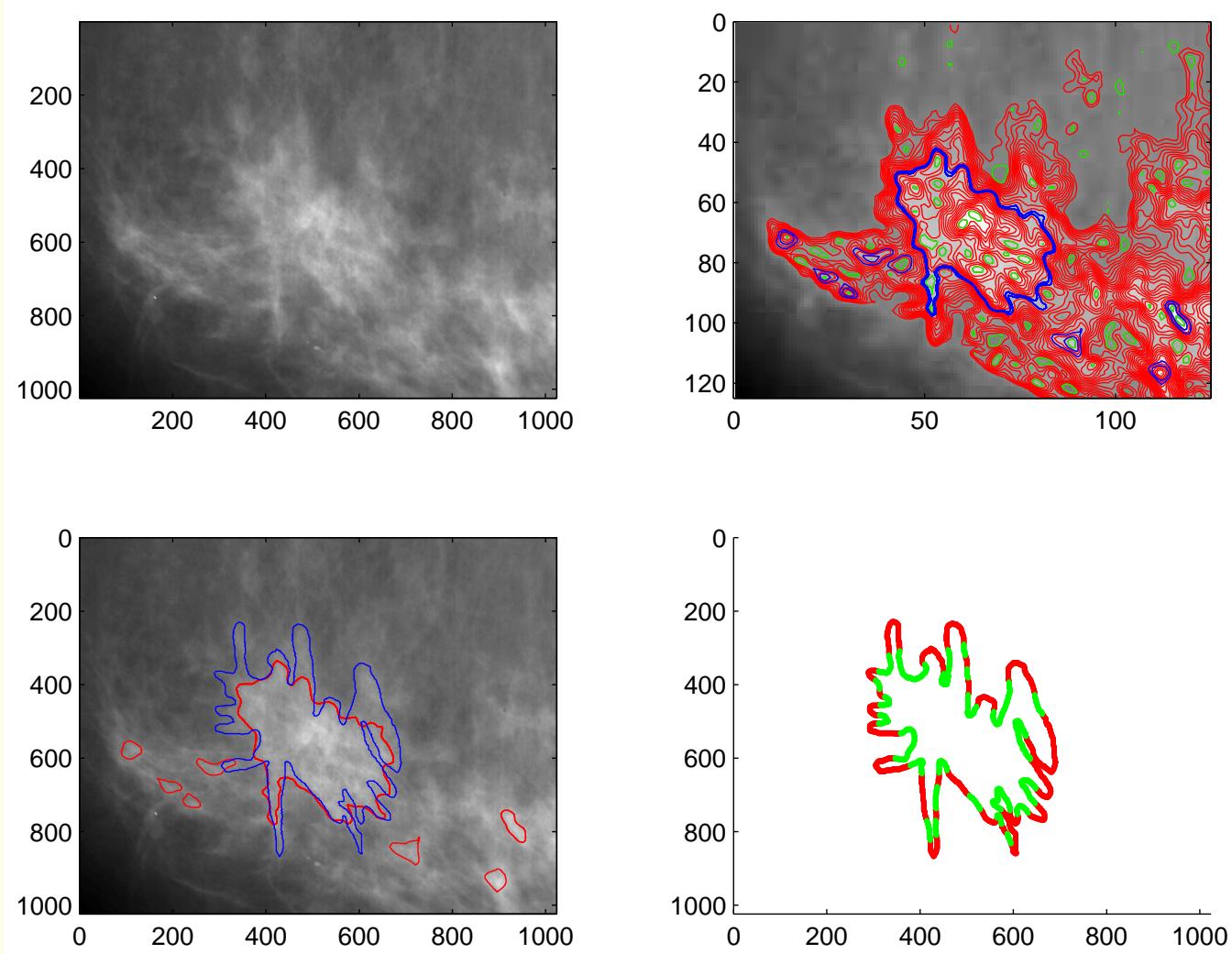


Illustration of various stages of
digital image processing and analysis.





1

The Nature of Biomedical Images

The human body is composed of many *systems*:

- visual system,
- cardiovascular system,
- musculo-skeletal system,
- central nervous system.



Each system is made up of several subsystems that carry on many *physiological processes*.

The visual system:

Performs the tasks of

- focusing visual or pictorial information on to the retina,
- transduction of the image information into neural signals,
- encoding and transmission of the neural signals to the visual cortex.

The visual cortex is responsible for interpretation of the image information.



The cardiac system:

Performs the important tasks of

- rhythmic pumping of blood through the arterial network of the body to facilitate the delivery of nutrients,
- pumping of blood through the pulmonary system for oxygenation of the blood itself.



The anatomical features of an organ or a system demonstrate characteristics that reflect

- the functional aspects of its processes, and
- the well-being or integrity of the system itself.



Physiological processes are complex phenomena, including

- neural or hormonal stimulation and control;
- inputs and outputs that could be in the form of physical material or information;
- action that could be mechanical, electrical, or biochemical.



Most physiological processes are accompanied by or manifest themselves as *signals* that reflect their nature and activities.

Such signals could be of many types, including

- biochemical in the form of hormones or neurotransmitters,
- electrical in the form of potential or current,
- physical in the form of pressure or temperature.



Diseases or defects in a physiological system cause alterations in its normal processes.

Pathological processes affect the performance, health, and general well-being of the system.

A pathological process is associated with signals and anatomical features that are different from the corresponding normal patterns.

With a good understanding of a system of interest, it becomes possible to observe the corresponding signals and features and assess the state of the system.



1.1 Body Temperature as an Image

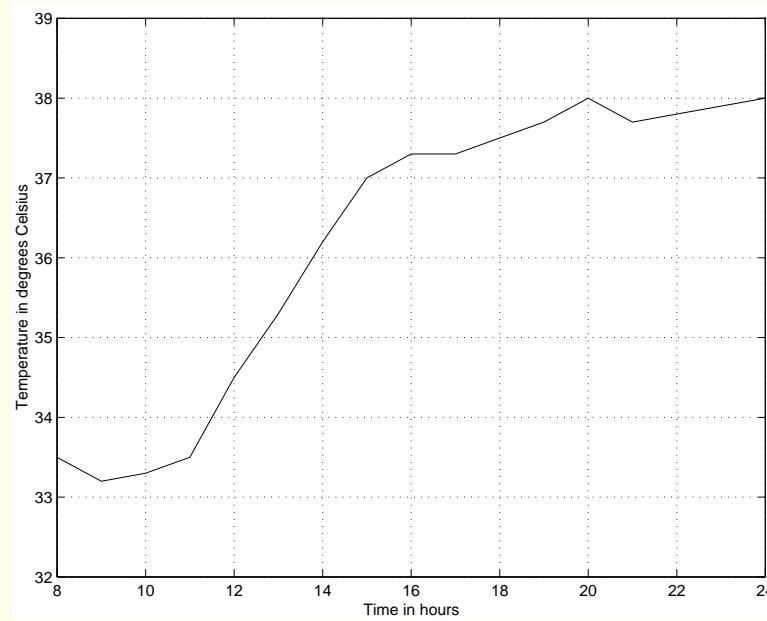
$33.5^{\circ}C$

(a)

Time (hours)	08	10	12	14	16	18	20	22	24
-----------------	----	----	----	----	----	----	----	----	----

Temperature ($^{\circ}C$)	33.5	33.3	34.5	36.2	37.3	37.5	38.0	37.8	38.0
--------------------------------	------	------	------	------	------	------	------	------	------

(b)



(c)

Figure 1.1: Measurements of the temperature of a patient presented as (a) a scalar with one temperature measurement f at a time instant t ; (b) an array $f(n)$ made up of several measurements at different instants of time; and (c) a signal $f(t)$ or $f(n)$. The horizontal axis of the plot represents time in hours; the vertical axis gives temperature in degrees Celsius. Data courtesy of Foothills Hospital, Calgary.



Infrared imaging:

Thermal sensors with wavelength in the range $3,000 - 5,000 \text{ nm}$ may be used to capture the heat radiated or emitted from a body or a part of a body as an image.

Thermal imaging has been investigated as a potential tool for the detection of breast cancer.

A tumor is expected to be more vascularized than its neighboring tissues, and hence could be at a slightly higher temperature.



The skin surface near the tumor may also demonstrate a relatively high temperature.

Temperature differences up to $2^{\circ}C$ have been measured between surface regions near breast tumors and neighboring tissues.

Thermography can help in the diagnosis of advanced cancer, but has limited success in the detection of early breast cancer.

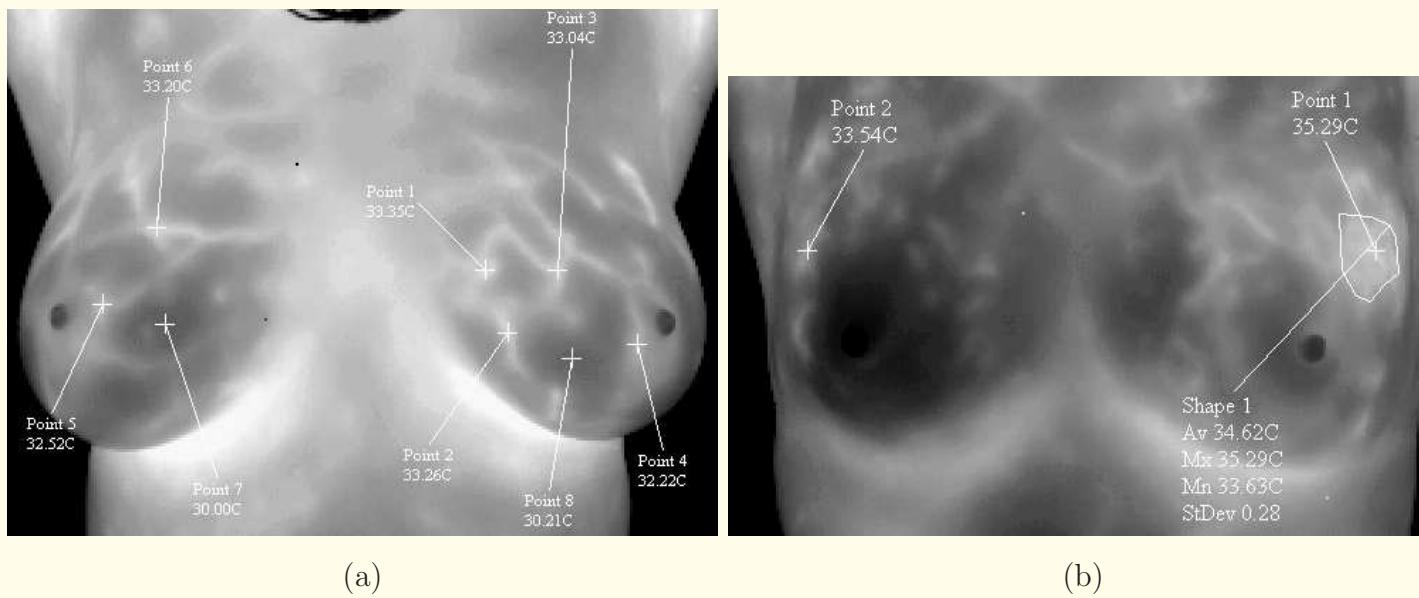


Figure 1.2: Body temperature as a 2D image $f(x, y)$ or $f(m, n)$. The images illustrate the distribution of surface temperature measured using an infrared camera operating in the $3,000 - 5,000 \text{ nm}$ wavelength range. (a) Image of a patient with pronounced vascular features and benign fibrocysts in the breasts. (b) Image of a patient with a malignant mass in the upper-outer quadrant of the left breast. Images courtesy of P. Hoekstra, III, Therma-Scan, Inc., Huntington Woods, MI.



Thermal images serve to illustrate an important distinction between two major categories of medical images:

- anatomical or physical images, and
- functional or physiological images.

A thermal image is a functional image.

An ordinary photograph obtained with reflected light is an anatomical or physical image.



1.2 Transillumination

Transillumination, diaphanography, and diaphanoscopy:

shining of visible light or near-infrared radiation through a part of the body, and viewing or imaging the transmitted radiation.



Application for the detection of breast cancer:

Nitrogen-rich compounds preferentially absorb (or attenuate) infrared radiation.

The fat and fibroglandular tissue in the mature breast contain much less nitrogen than malignant tissues.

The hemoglobin in blood has a high nitrogen content, and tumors are more vascularized than normal tissues.

For these reasons, breast cancer appears as a relatively dark region in a transilluminated image.



The effectiveness of transillumination is limited by scatter and ineffective penetration of light through a large organ.

Transillumination has been found to be useful in differentiating between cystic (fluid-filled) and solid lesions.

However, the technique has had limited success in distinguishing malignant tumors from benign masses.



1.3 Light Microscopy

Useful magnification of up to $\times 1,000$ may be obtained via light microscopy by the use of combinations of lenses.

The resolution of light microscopy is reduced by the following:

- **Diffraction:** The bending of light at edges causes blurring; the image of a pinhole appears as a blurred disc known as the Airy disc.
- **Astigmatism:** Due to nonuniformities in lenses, a point may appear as an ellipse.



- **Chromatic aberration:** Electromagnetic (EM) waves of different wavelength or energy that compose the ordinarily used white light converge at different focal planes, thereby causing enlargement of the focal point.

This effect may be prevented by using monochromatic light.

- **Spherical aberration:** The rays of light arriving at the periphery of a lens are refracted more than the rays along the axis of the lens.

This causes the rays from the periphery and the axis not to arrive at a common focal point, thereby resulting in blurring.

The effect may be reduced by using a small aperture.



- **Geometric distortion:** Poorly crafted lenses may cause geometric distortion such as the pin-cushion effect and barrel distortion.

Whereas the best resolution achievable by the human eye is of the order of $0.1 - 0.2 \text{ mm}$, light microscopes can provide resolving power up to about $0.2 \text{ } \mu\text{m}$.



Example: A single ventricular myocyte (of a rabbit) in its relaxed state as seen through a light microscope at a magnification of about $\times 600$.

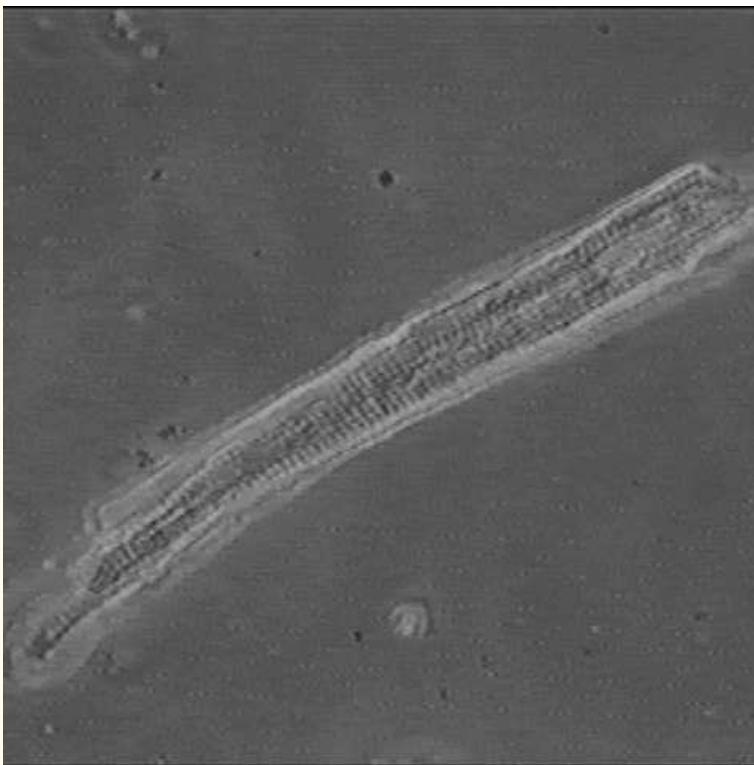


Figure 1.3: A single ventricular myocyte (of a rabbit) in its relaxed state as seen through a light microscope at a magnification of about $\times 600$. The width (thickness) of the myocyte is approximately $15 \mu\text{m}$. Image courtesy of R. Clark, Department of Physiology and Biophysics, University of Calgary.



Example: Figure 1.4 shows images of three-week-old scar tissue and forty-week-old healed tissue samples from rabbit ligaments at a magnification of about $\times 300$.

The images demonstrate the alignment patterns of the nuclei of fibroblasts (stained to appear as the dark objects in the images).

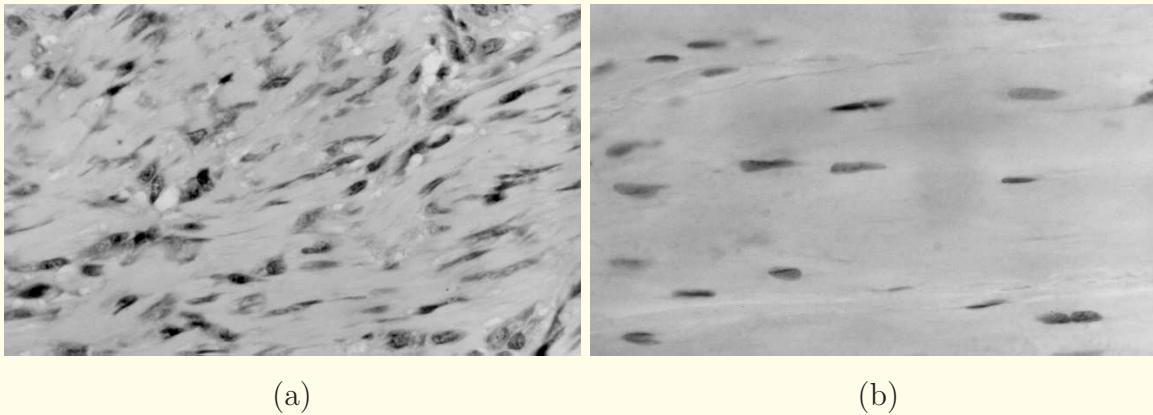


Figure 1.4: (a) Three-week-old scar tissue sample, and (b) forty-week-old healed tissue sample from rabbit medial collateral ligaments. Images courtesy of C.B. Frank, Department of Surgery, University of Calgary.



1.4 Electron Microscopy

Accelerated electrons possess EM wave properties, with the wavelength λ given by

$$\lambda = \frac{h}{mv},$$

where h is Planck's constant, m is the mass of the electron, and v is the electron's velocity.



This relationship reduces to

$$\lambda = \frac{1.23}{\sqrt{V}},$$

where V is the accelerating voltage.

At 60 kV , an electron beam has an effective wavelength of about 0.005 nm , and a resolving power limit of about 0.003 nm .



A low kV provides high contrast but low resolution.

A high kV provides high resolution due to smaller wavelength but low contrast due to higher penetrating power.

A high- kV beam causes less damage to the specimen as the faster electrons pass through the specimen in less time than with a low- kV beam.

Electron microscopes can provide useful magnification of the order of $\times 10^6$, and may be used to reveal the ultrastructure of biological tissues.

Electron microscopy typically requires the specimen to be fixed, dehydrated, dried, mounted, and coated with a metal.



Transmission electron microscopy:

A transmission electron microscope (TEM) consists of

- a high-voltage electron beam generator,
- a series of EM lenses,
- a specimen holding and changing system,
- a screen-film holder,

all enclosed in vacuum.



In TEM, the electron beam passes through the specimen, is affected in a manner similar to light.

The resulting image is captured through a screen-film combination or viewed via a phosphorescent viewing screen.



Example: Figure 1.5 shows TEM images of collagen fibers (in cross-section) in rabbit ligament samples.

Scar samples have been observed to have an almost uniform distribution of fiber diameter in the range $60 - 70\text{ nm}$.

Normal samples have an average diameter of about 150 nm over a broader distribution.

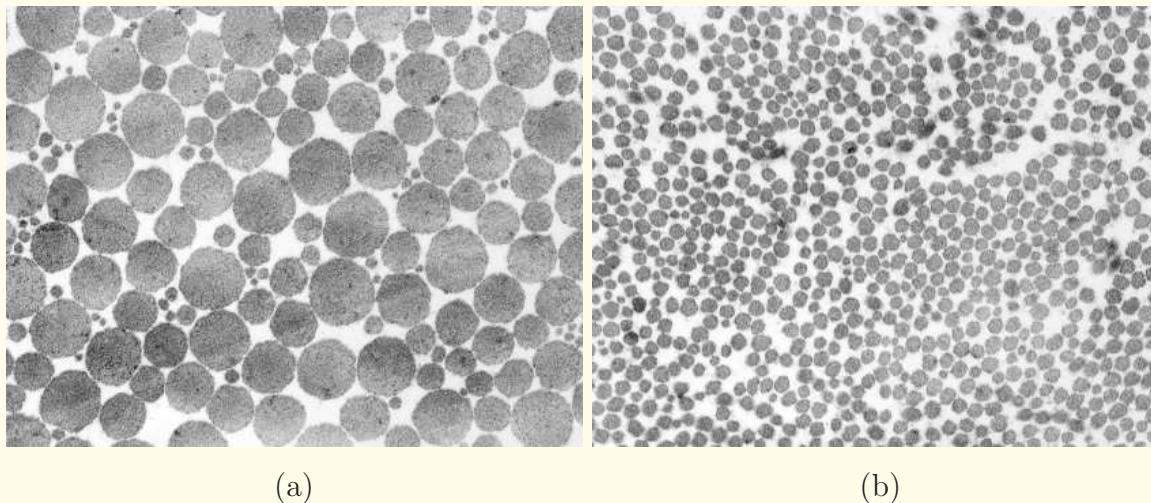


Figure 1.5: TEM images of collagen fibers in rabbit ligament samples at a magnification of approximately $\times 30,000$. (a) Normal and (b) scar tissue. Images courtesy of C.B. Frank, Department of Surgery, University of Calgary.



Example: In patients with hematuria, the glomerular basement membrane of capillaries in the kidney is thinner ($< 200 \text{ nm}$) than the normal thickness of the order of 300 nm .

Investigation of this feature requires needle-core biopsy of the kidney and TEM imaging.

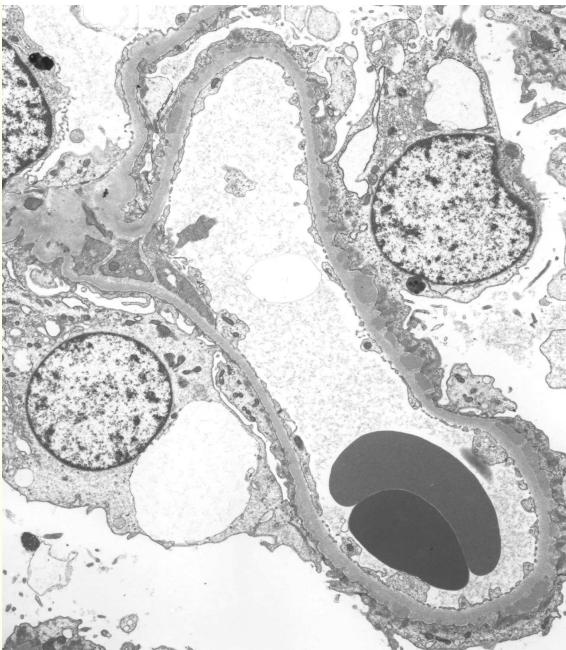
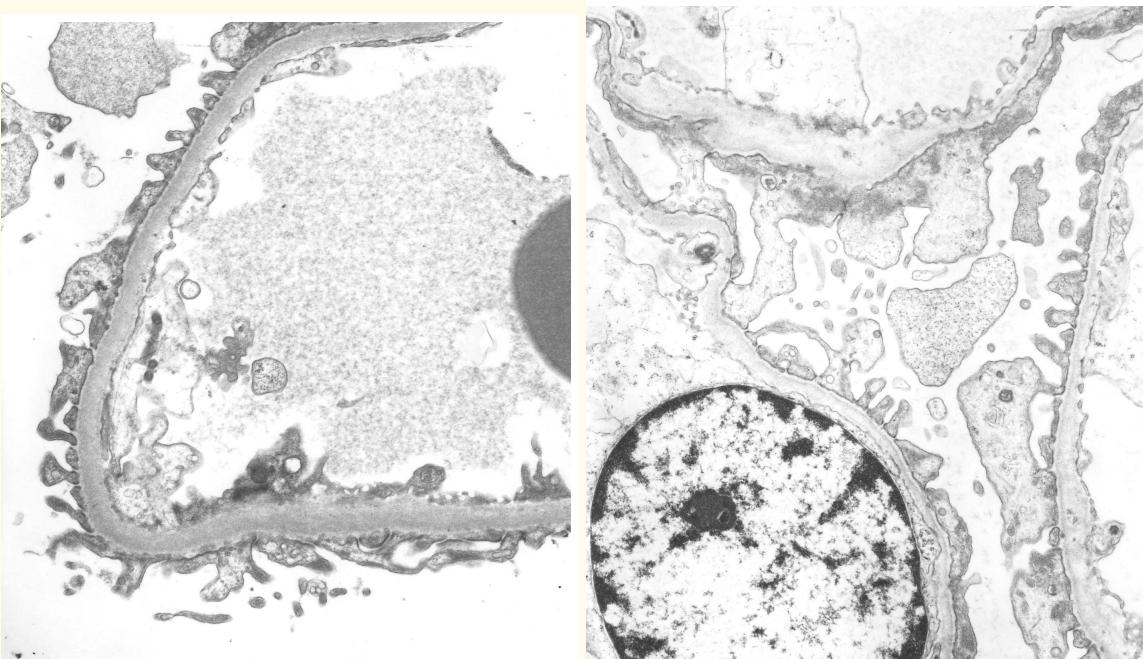


Figure 1.6: TEM image of a kidney biopsy sample at a magnification of approximately $\times 3,500$. The image shows the complete cross-section of a capillary with normal membrane thickness. Image courtesy of H. Benediktsson, Department of Pathology and Laboratory Medicine, University of Calgary.



(a)

(b)

Figure 1.7: TEM images of kidney biopsy samples at a magnification of approximately $\times 8,000$. (a) The sample shows normal capillary membrane thickness. (b) The sample shows reduced and varying membrane thickness. Images courtesy of H. Benediktsson, Department of Pathology and Laboratory Medicine, University of Calgary.



Scanning electron microscopy:

A scanning electron microscope (SEM) is similar to a TEM in many ways, but uses a finely focused electron beam with a diameter of the order of 2 nm to scan the surface of the specimen in a raster pattern.

The secondary electrons that are emitted from the surface of the sample are detected and amplified through a photomultiplier tube (PMT), and used to form an image on a CRT.

An SEM may be operated in different modes to detect a variety of signals emitted from the sample, and may be used to obtain images with a depth of field of several mm .



Example: Figure 1.8 illustrates SEM images of collagen fibers in rabbit ligament samples (freeze-fractured surfaces).

The images are useful in analyzing the angular distribution of fibers and the realignment process during healing after injury.

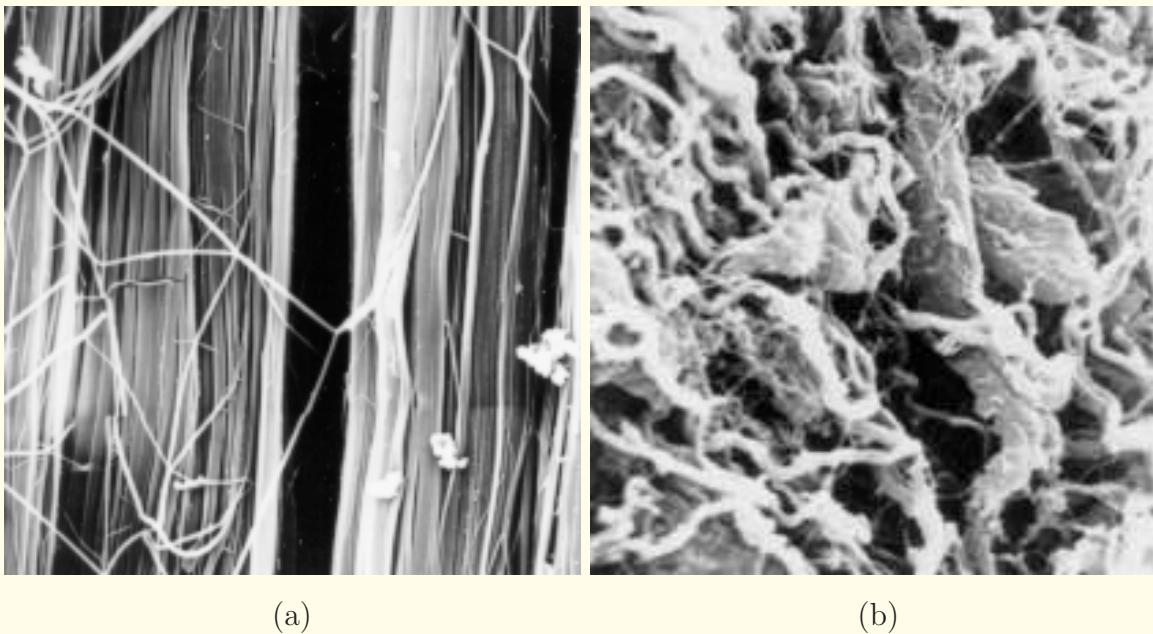


Figure 1.8: SEM images of collagen fibers in rabbit ligament samples at a magnification of approximately $\times 4,000$.
(a) Normal and (b) scar tissue.



1.5 X-ray Imaging

Planar X-ray imaging or radiography: a 2D projection (shadow or silhouette) of a 3D body is produced on film by irradiating the body with X-ray photons.

Each ray of X-ray photons is attenuated by a factor depending upon the integral of the linear attenuation coefficient along the path of the ray, and produces a corresponding gray level (signal) at the point hit on the film or the detecting device used.



Beer's law or Beer–Lambert law:

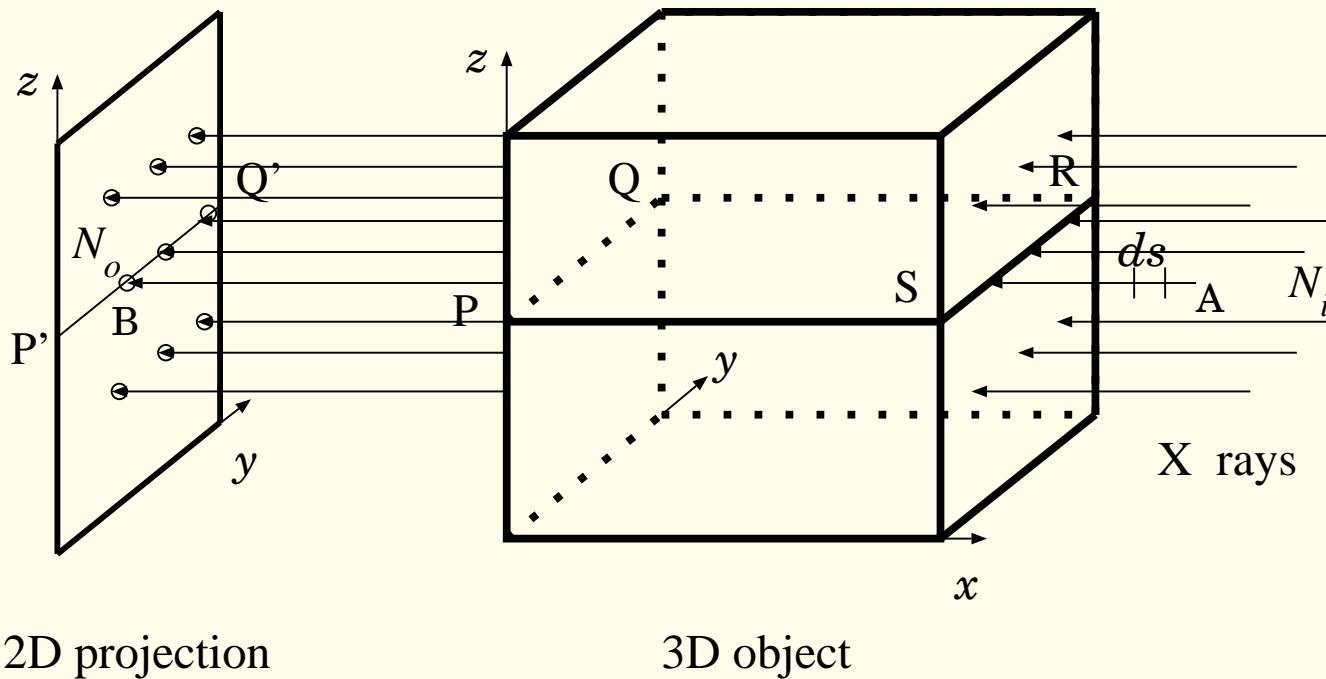
Consider the ray path marked as AB in Figure 1.9.

Let N_i denote the number of X-ray photons incident upon the body being imaged, within a specified time interval.

Let N_o be the number of photons exiting the body.

The mutually parallel rays within the plane PQRS are represented by the coordinates (t, s) that are at an angle θ with respect to the (x, y) coordinates, with the s axis being parallel to the rays.

$$s = -x \sin \theta + y \cos \theta.$$



2D projection

3D object

Figure 1.9: An X-ray image or a typical radiograph is a 2D projection or planar image of a 3D object. The entire object is irradiated with X rays. The projection of a 2D cross-sectional plane $PQRS$ of the object is a 1D profile $P'Q'$ of the 2D planar image.



$$N_o = N_i \exp \left[- \int_{rayAB} \mu(x, y) ds \right], \quad (1.1)$$

or

$$\int_{rayAB} \mu(x, y) ds = \ln \left(\frac{N_i}{N_o} \right). \quad (1.2)$$

The use of monochromatic or monoenergetic X rays is assumed.

N_i, N_o : Poisson variables; large values assumed.

$\mu(x, y)$: linear attenuation coefficient at (x, y) in the sectional plane PQRS.

$\mu(x, y)$ depends upon the density of the object and the frequency (wavelength or energy) of the radiation used.



A measurement of the exiting X rays (that is, N_o , and N_i for reference) thus gives us only an integral of $\mu(x, y)$ over the ray.

The internal details of the body along the ray path are compressed onto a single point on the film or a single measurement.

The radiographic image produced is a 2D planar image of the 3D object, where the internal details are superimposed.



When the rays are parallel to the x axis, we have

$$\theta = 90^\circ, s = -x, ds = -dx,$$

and the planar image is given by

$$g(y, z) = \int -\mu(x, y, z) dx. \quad (1.3)$$



Screen-film detector:

The X rays exiting from the body strike a fluorescent (phosphor) screen made of compounds of rare-earth elements such as lanthanum oxybromide or gadolinium oxysulfide.

The X-ray photons are converted into visible-light photons.

A light-sensitive film in contact with the screen (in a light-tight cassette) records the result.

The film contains a layer of silver-halide emulsion with a thickness of about $10 \mu\text{m}$.

The exposure or blackening of the film depends upon the number of light photons that reach the film.



A thick screen provides a high efficiency of conversion of X rays to light, but causes loss of resolution due to blurring.

The typical thickness of the phosphor layer in screens is in the range $40 - 100 \mu\text{m}$.

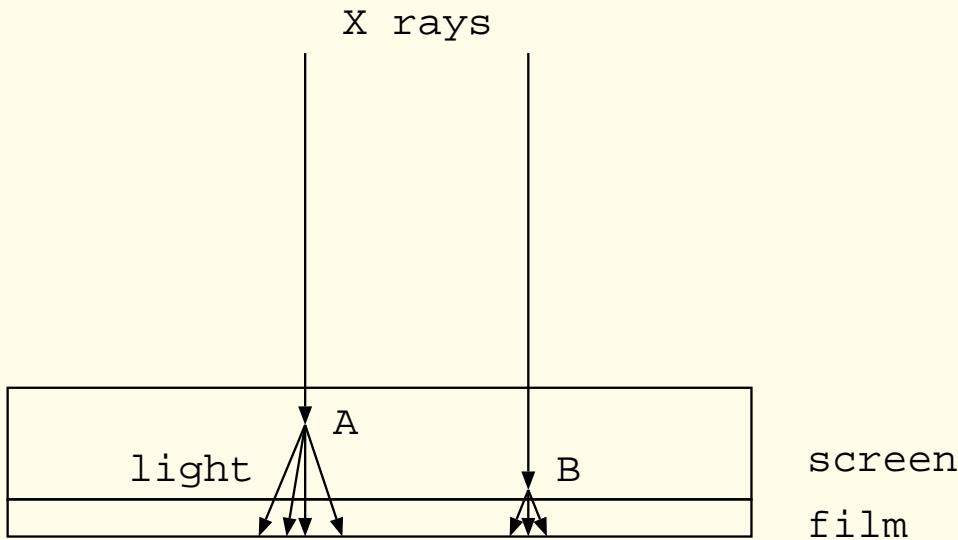


Figure 1.10: Blur caused by a thick screen. Light emanating from point A in the screen is spread over a larger area on the film than that from point B.



A fluoroscopy system uses an image intensifier and a video camera in place of the film to capture the image and display it on a monitor as a movie or video.

Images are acquired at a rate of $2 - 8 \text{ frames/s (fps)}$, with the X-ray beam pulsed at $30 - 100 \text{ ms}$ per frame.

In computed radiography (CR), a photo-stimulable phosphor plate (made of europium-activated barium fluorohalide) is used instead of film to capture and temporarily hold the image.

The latent image is scanned using a laser and digitized.

In digital radiography (DR), the film or the entire screen-film combination is replaced with solid-state electronic detectors.



Examples: Figures 1.11 (a) and (b) show the posterior-anterior (PA, that is, back-to-front) and lateral (side-to-side) X-ray images of the chest of a patient.



(a)



(b)

Figure 1.11: (a) Posterior-anterior and (b) lateral chest X-ray images of a patient. Images courtesy of Foothills Hospital, Calgary.



Physical and technical considerations

- **Target and focal spot:** An electron beam with energy in the range of $20 - 140 \text{ keV}$ is used to produce X rays for diagnostic imaging.

Typical target materials used: tungsten and molybdenum.

Focal spot: area of the target struck by the electron beam to generate X rays.

Nominal focal spot: diameter in mm as observed in the imaging plane (on the film).

A small focal spot is desired in order to obtain a sharp image, especially in magnification imaging.

Typical focal spot sizes in radiography: $0.1 - 2 \text{ mm}$.

A focal spot size of $0.1 - 0.3 \text{ mm}$ is desired in mammography.



- **Energy:** The penetrating capability of an X-ray beam is mainly determined by the accelerating voltage applied to the electron beam that impinges the target in the X-ray generator. Indicator of penetrating capability (the “energy” of the X-ray beam): kVp , kilo-volt-peak.
Higher kVp : more penetrating X-ray beam.



The actual unit of energy of an X-ray photon is the electron volt or eV , which is the energy gained by an electron when a potential of 1 V is applied to it.

The kVp measure relates to the highest possible X-ray photon energy that may be achieved at the voltage used.

Low-energy X-ray photons are absorbed at or near the skin surface, and do not contribute to the image.

In order to prevent unwanted radiation, a filter is used at the X-ray source to absorb low-energy X rays.

Typical filter materials: aluminum and molybdenum.



Imaging of soft-tissue organs such as the breast is performed with low-energy X rays in the range of $25 - 32 \text{ kVp}$.

The use of a higher kVp would result in low differential attenuation and poor tissue-detail visibility or contrast.

A few other energy levels used in projection radiography are:

- abdomen: $60 - 100 \text{ kVp}$;
- chest: $80 - 120 \text{ kVp}$;
- skull: $70 - 90 \text{ kVp}$.

The kVp to be used depends upon the distance between the X-ray source and the patient, the size (thickness) of the patient, the type of grid used, and several other factors.



- **Exposure:** For a given tube voltage (kVp), the total number of X-ray photons released at the source is related to the product of the tube current (mA) and the exposure time (s), together expressed as the product mAs .

For a given body being imaged, the number of photons that arrive at the film is related to the mAs quantity.

A low mAs results in an under-exposed film (faint or light image), whereas a high mAs results in an over-exposed or dark image (as well as increased X-ray dose to the patient).



Typical exposure values: $2 - 120 \text{ mAs}$.

Most imaging systems determine automatically the required exposure for a given mode of imaging, patient size, and kVp .

Some systems use an initial exposure of the order of 5 ms to estimate the penetration of the X rays through the body being imaged, and then determine the required exposure.



▪ **Beam hardening:** The X rays used in radiographic imaging are typically not monoenergetic; they possess X-ray photons over a certain band of frequencies or EM energy levels.

As the X rays propagate through a body, the lower-energy photons get absorbed preferentially, depending upon the length of the ray path through the body and the attenuation characteristics of the tissues along the path.

The X rays that pass through the object at longer distances from the source will possess relatively fewer photons at lower-energy levels than at the point of entry (and hence a relatively higher concentration of higher-energy photons).



This is known as beam hardening: leads to incorrect estimation of the attenuation coefficient in CT imaging.

The effect of beam hardening may be reduced by prefiltering or prehardening the X-ray beam and narrowing its spectrum.

The use of monoenergetic X rays from a synchrotron or a laser obviates this problem.



▪ **Scatter and the use of grids:** As an X-ray beam propagates through a body, photons are lost due to absorption and scattering at each point in the body.

The angle of the scattered photon is a random variable.

A scattered photon contributes to noise at the point where it strikes the detector.

Scattering results in the loss of contrast.

The noise effect of the scattered radiation is significant in gamma-ray emission imaging.

The effect of scatter may be reduced by the use of grids, collimation, or energy discrimination.

Scattered (secondary) photons usually have lower energy levels than the primary photons.



- **Grid:** Array of mutually parallel X-ray absorbing strips if the X rays are in a parallel beam (as in chest imaging), or converging toward the X-ray source in the case of a diverging beam (as in breast imaging).

Lattice or honeycomb grids with parallel strips in criss-cross patterns are also used in mammography.

X-ray photons that arrive via a path that is not aligned with the grids will be stopped from reaching the detector.

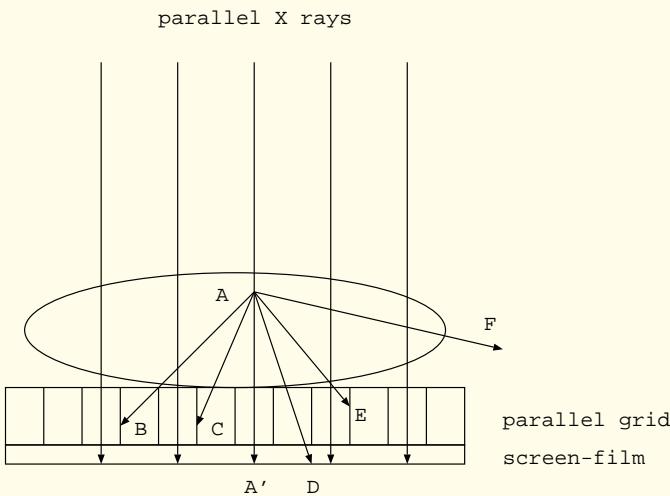


Figure 1.12: Use of parallel grids to reduce scatter. X rays that are parallel to the grids reach the film; for example, line AA'. Scattered rays AB, AC, and AE have been blocked by the grids; however, the scattered ray AD has reached the film in the illustration.



A typical grid contains thin strips of lead or aluminum with a strip density of $25 - 80 \text{ lines/cm}$ and a grid height to strip width ratio in the range of $5 : 1$ to $12 : 1$.

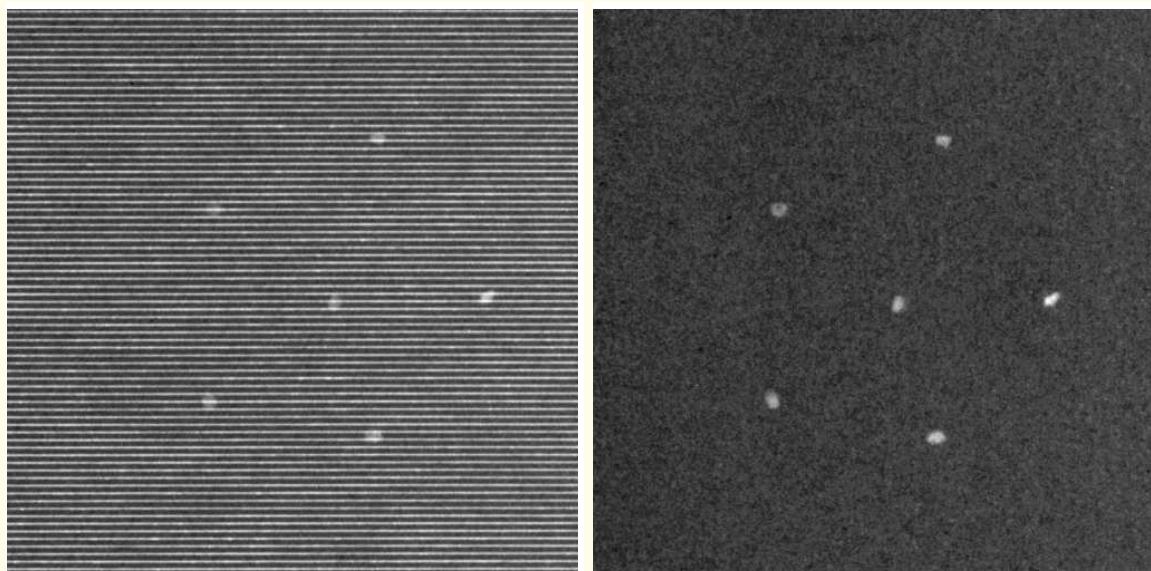
The space between the grids is filled with low-attenuation material such as wood.

A stationary grid produces a line pattern that is superimposed upon the image.

Grid artifact is prevented in a reciprocating grid, where the grid is moved about 20 grid spacings during exposure: the movement smears the grid shadow and renders it invisible.

Low levels of grid artifact may appear if the bucky does not move at a uniform pace or starts moving late or ends movement early with respect to the X-ray exposure interval.

Disadvantages: double radiation dose, reduced contrast.



(a)

(b)

Figure 1.13: X-ray images of a part of a phantom: (a) with, and (b) without grid artifact. Image courtesy of L.J. Hahn, Foothills Hospital, Calgary.



Figure 1.14: X-ray image of the American College of Radiology (ACR) phantom for mammography. The pixel-value range [117, 210] has been linearly stretched to the display range [0, 255] to show the details. Image courtesy of S. Bright, Sunnybrook & Women's College Health Sciences Centre, Toronto, ON, Canada.



- **Photon detection noise:**

Interaction between an X-ray beam and a detector:

Photons lost due to scatter and absorption.

Some photons may pass through unaffected (or undetected).

The small size of the detectors in DR and CT imaging reduces their detection efficiency.

Scattered and undetected photons cause noise.



- **Ray stopping by heavy implants:**

Extremely heavy parts or components, such as metal screws or pins in bones and surgical clips that are nearly X-ray-opaque and entirely stop the incoming X-ray photons, can completely block an X-ray beam.

No photons would be detected at the corresponding point of exit from the body.

The attenuation coefficient for the corresponding path would be indefinite, or within the computational context, infinity.

A reconstruction algorithm would not be able to redistribute the attenuation values over the points along the corresponding ray path in the reconstructed image.

This leads to streaking artifacts in CT images.



Special techniques for enhanced X-ray imaging:

- digital subtraction angiography (DSA).
- dual-energy imaging.



1.5.1 *Breast cancer and mammography*

Cancer is caused when a single cell or a group of cells escapes from the usual controls that regulate cellular growth, and begins to multiply and spread.

This activity results in a mass, tumor, or neoplasm.

Many masses are benign; that is, the abnormal growth is restricted to a single, circumscribed, expanding mass of cells.

Some tumors are malignant: the abnormal growth invades the surrounding tissues and may spread, or metastasize, to distant areas of the body.



Benign masses may lead to complications.

Malignant tumors are serious: **cancer**.

The majority of breast tumors will have metastasized before reaching a palpable size.

Mammography has gained recognition as the single most successful technique for the detection of early, clinically occult breast cancer.



Mammograms are analyzed by radiologists specialized in mammography.

A normal mammogram typically depicts converging patterns of fibroglandular tissues and vessels.

Any feature that causes a departure from or distortion with reference to the normal pattern is viewed with suspicion and analyzed with extra attention.

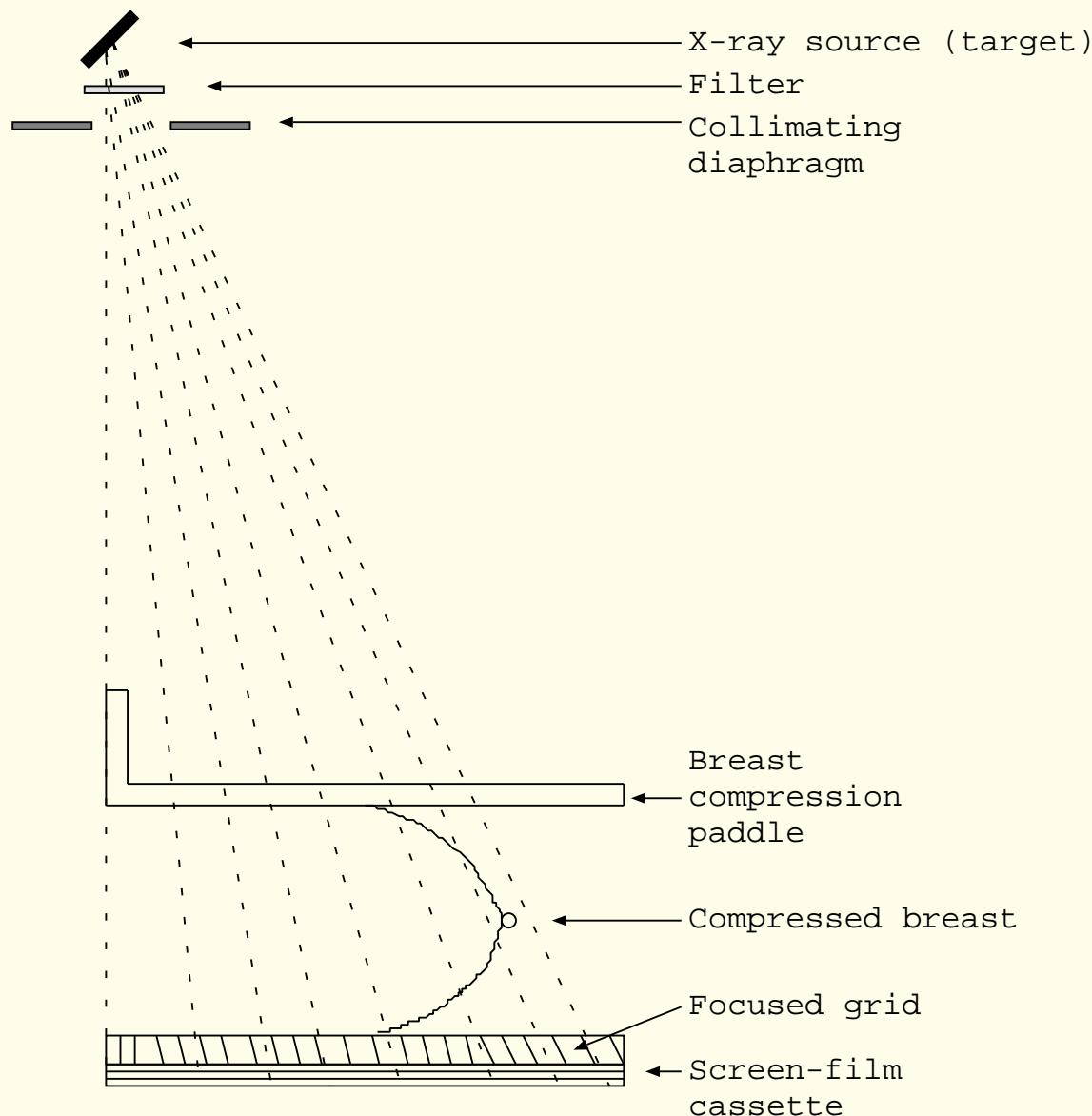
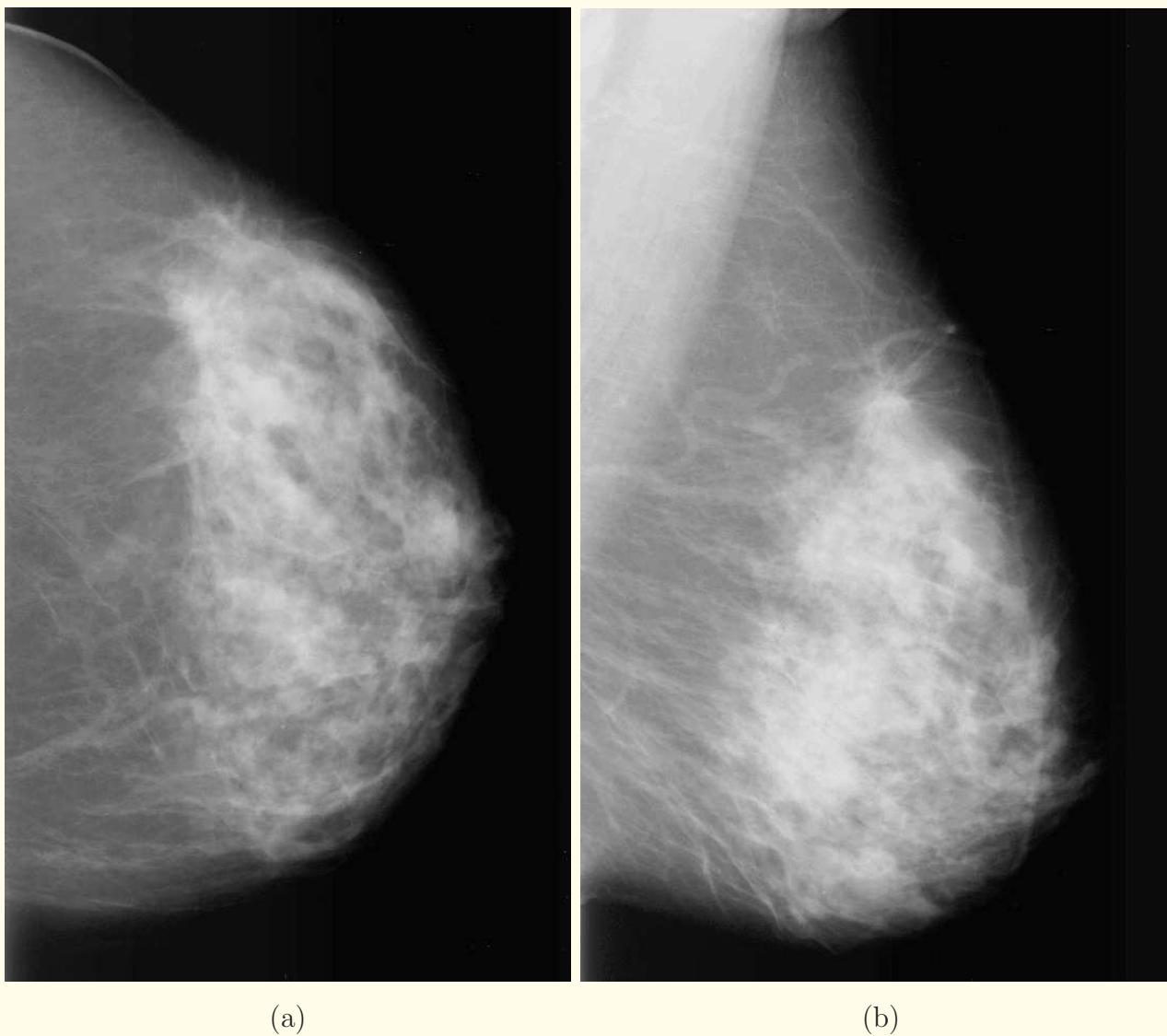


Figure 1.15: A typical mammography setup.



(a)

(b)

Figure 1.16: (a) Cranio-caudal (CC) and (b) medio-lateral oblique (MLO) mammograms of the same breast of a subject. The MLO view demonstrates architectural distortion due to a spiculated tumor near the upper right-hand corner edge. Images courtesy of Screen Test — Alberta Program for the Early Detection of Breast Cancer.



Important signs of abnormalities and cancer:

- calcifications,
- masses,
- localized increase in density,
- asymmetry between the left and right breast images, and
- architectural distortion.



1.6 Tomography

Problem: visualizing the details of the interior of the human body or other objects.

Laminagraphy, planigraphy, or “classical” tomography used synchronous movement of the X-ray source and film in such a way as to produce a relatively sharp image of a single focal plane of the object, with the images of all other planes being blurred.

The smearing of information from the other planes of the object causes loss of contrast in the plane of interest.

CT imaging made film-based tomography obsolete.

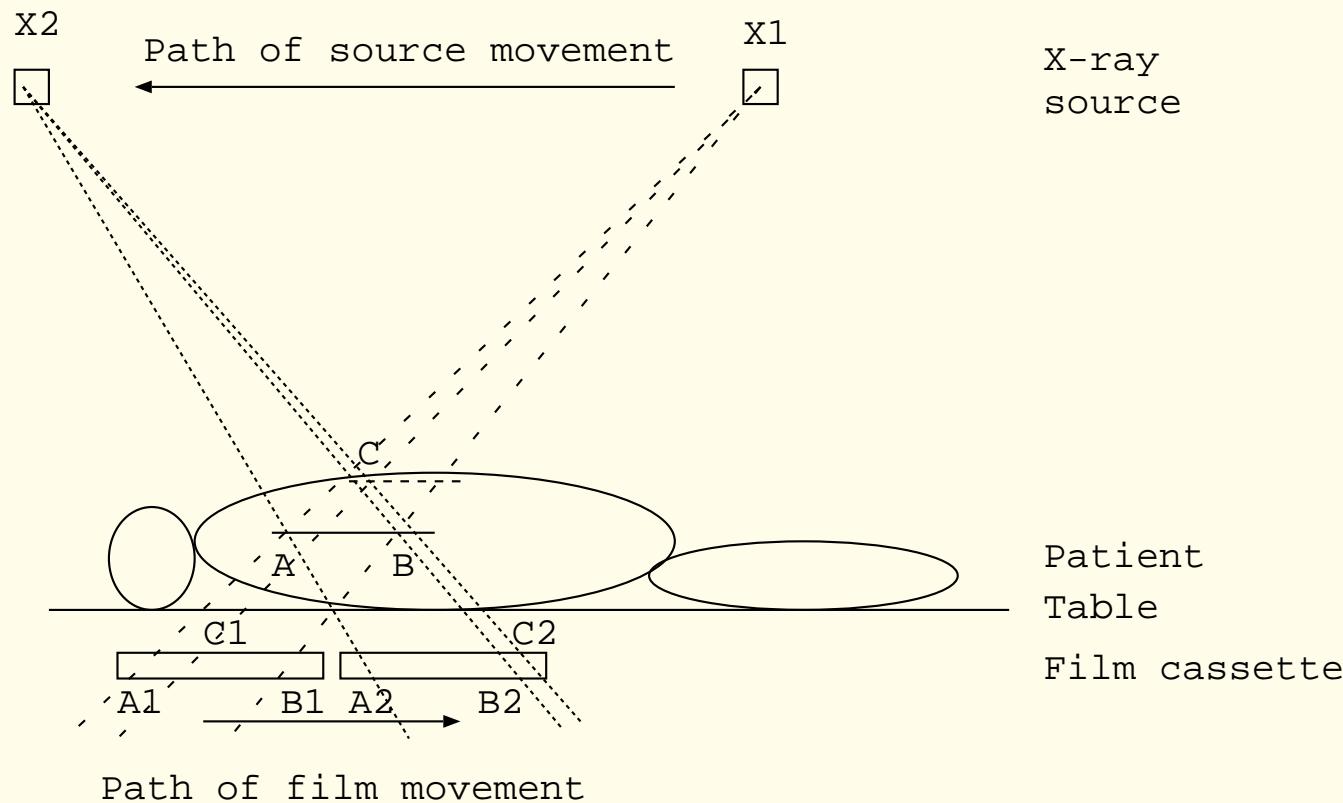


Figure 1.17: Synchronized movement of the X-ray source and film to obtain a tomographic image of the focal plane indicated as AB.



Figure 1.18: Tomographic image of a patient in a longitudinal (coronal) plane through the chest. Reproduced with permission from R.A. Robb, "X-ray computed tomography: An engineering synthesis of multiscientific principles", *CRC Critical Reviews in Biomedical Engineering*, 7:264–333, March 1982. © CRC Press.

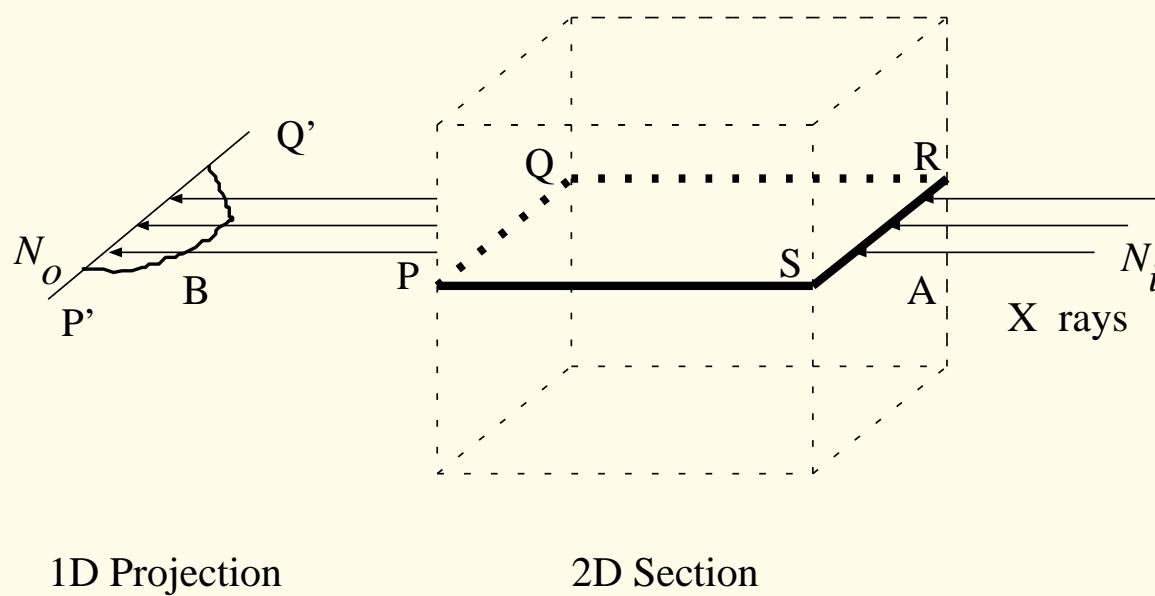


Computed tomography: The technique of CT imaging was developed during the late 1960s and the early 1970s, producing images of cross-sections of the human head and body as never seen before (*noninvasively and nondestructively!*).

In the simplest form of CT imaging, only the desired cross-sectional plane of the body is irradiated using a finely collimated ray of X-ray photons.

Ray integrals are measured at many positions and angles around the body, scanning the body in the process.

The principle of image reconstruction from projections, is then used to compute an image of a section of the body: hence the name *computed* tomography.



1D Projection

2D Section

Figure 1.19: In the basic form of CT imaging, only the cross-sectional plane of interest is irradiated with X rays. The projection of the 2D cross-sectional plane PQRS of the object is the 1D profile P'Q' shown. Compare this case with the planar imaging case illustrated in Figure 1.9. Reproduced, with permission, from R.M. Rangayyan and A. Kantzas, “Image reconstruction”, *Wiley Encyclopedia of Electrical and Electronics Engineering*, Supplement 1, Editor: John G. Webster, Wiley, New York, NY, pp 249–268, 2000. © This material is used by permission of John Wiley & Sons, Inc.

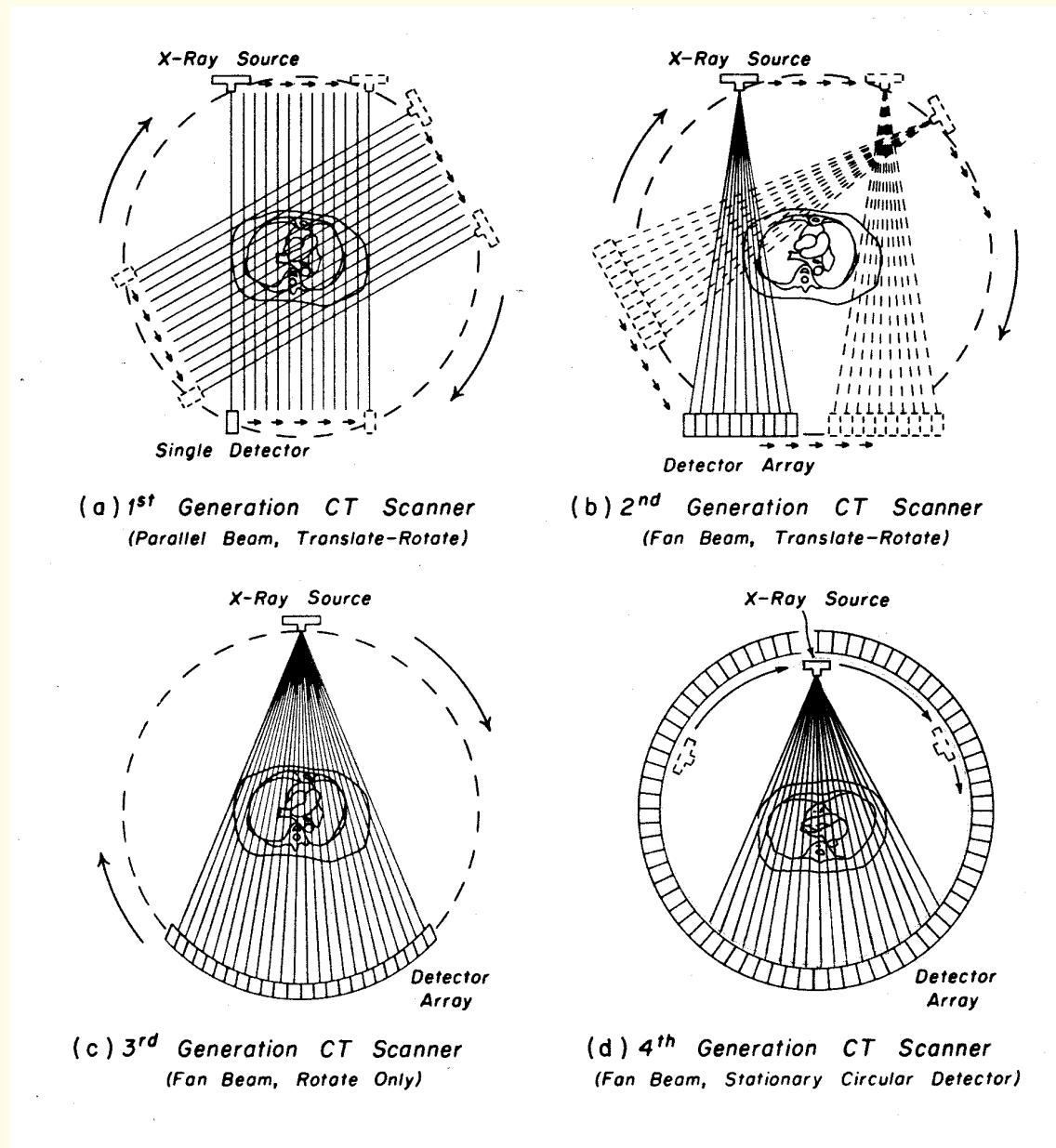


Figure 1.20: (a) Translate-rotate scanning geometry for parallel-ray projections; (b) translate-rotate scanning geometry with a small fan-beam detector array; (c) rotate-only scanning geometry for fan-beam projections; (d) rotate-only scanning geometry for fan-beam projections using a ring of detectors. Reproduced with permission from R.A. Robb, "X-ray computed tomography: An engineering synthesis of multiscientific principles", *CRC Critical Reviews in Biomedical Engineering*, 7:264–333, March 1982. © CRC Press.

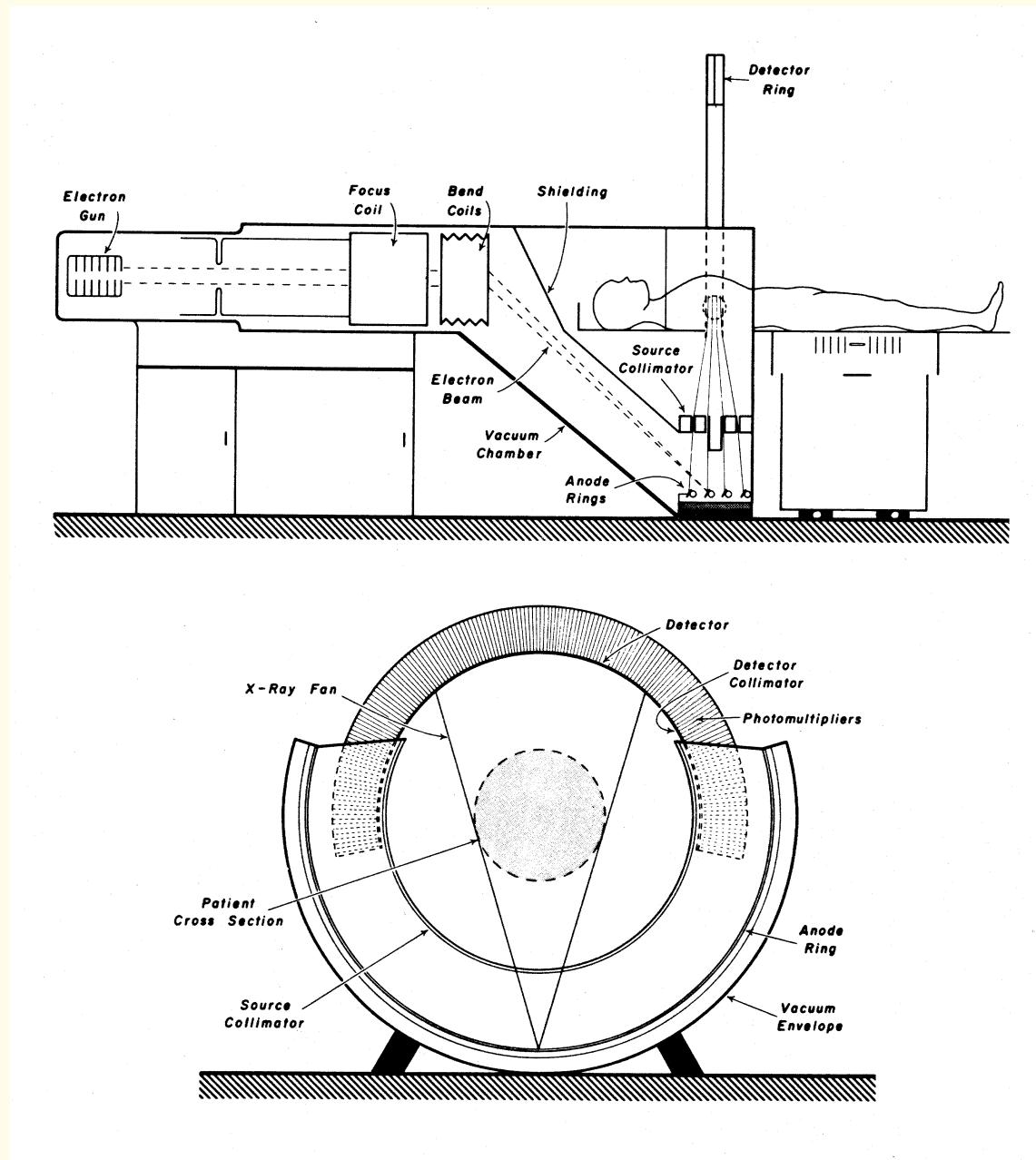


Figure 1.21: Electronic steering of an X-ray beam for motion-free scanning and CT imaging. Reproduced with permission from D.P. Boyd, R.G. Gould, J.R. Quinn, and R. Sparks, "Proposed dynamic cardiac 3-D densitometer for early detection and evaluation of heart disease", *IEEE Transactions on Nuclear Science*, 26(2):2724–2727, 1979. © IEEE.



Figure 1.22: CT image of a patient showing the details in a cross-section through the head (brain). Image courtesy of Foothills Hospital, Calgary.

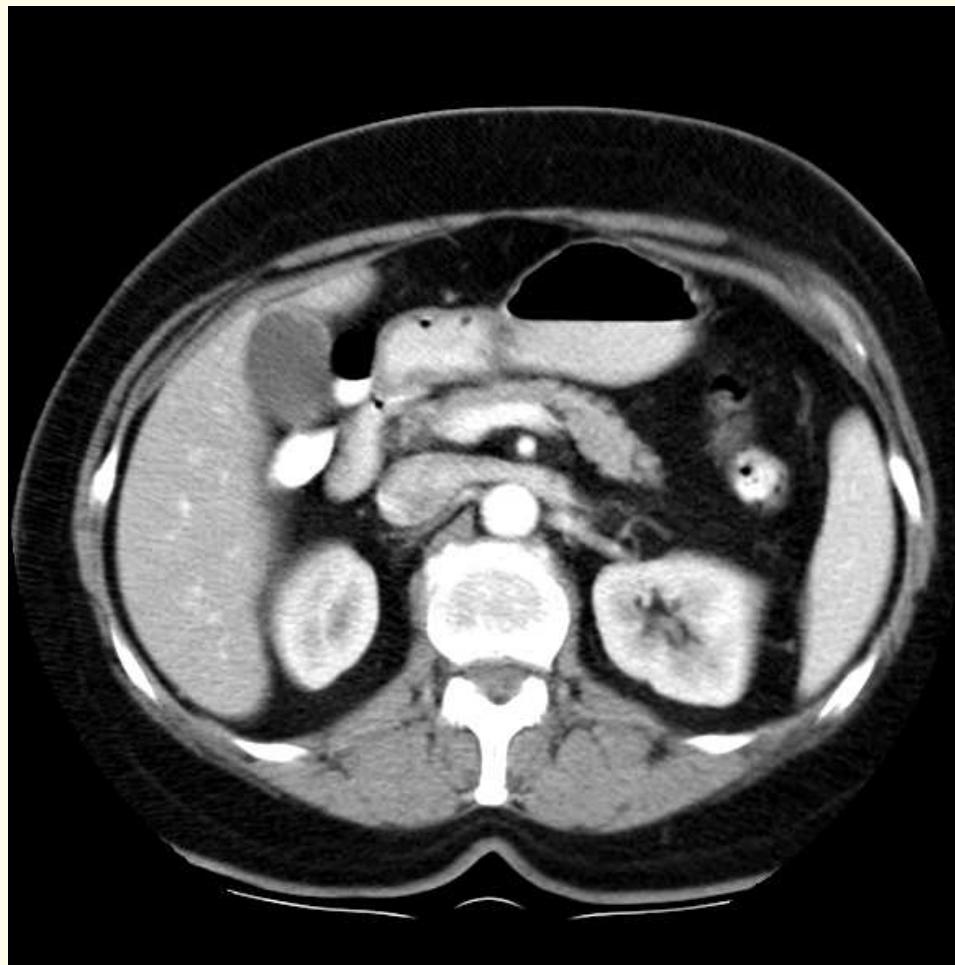


Figure 1.23: CT image of a patient showing the details in a cross-section through the abdomen. Image courtesy of Foothills Hospital, Calgary.



(a)



(b)

Figure 1.24: CT image of a patient scaled to (a) show the details of the lungs; and (b) display the mediastinum in detail — the details of the lungs are not visible in this rendition. Images courtesy of Alberta Children's Hospital, Calgary.



1.7 Nuclear Medicine Imaging

In nuclear medicine imaging, a small quantity of a radiopharmaceutical is administered into the body orally, by intravenous injection, or by inhalation.

The radiopharmaceutical is designed so as to be absorbed by and localized in a specific organ of interest.

The gamma-ray photons emitted as a result of radioactive decay of the radiopharmaceutical are used to form an image that represents the distribution of radioactivity in the organ.

Nuclear medicine imaging is used to map physiological function such as perfusion and ventilation of the lungs, and blood supply to the musculature of the heart, liver, spleen, and thyroid gland.



Whereas X-ray images provide information related to density and may be used to detect altered anatomy, nuclear medicine imaging helps in examining altered physiological (or pathological) functioning of specific organs in a body.

Commonly used isotopes in nuclear medicine imaging:

- Technetium as ^{99m}Tc — gamma-ray photons at 140 keV .
- Thallium as ^{201}Tl at 70 keV or 167 keV .
- Iodine as ^{131}I (for thyroid imaging).



The scintillation gamma camera or the Anger camera uses a large thallium-activated sodium iodide [$\text{NaI}(Tl)$] detector, typically *40 cm* in diameter and *10 mm* in thickness.

The gamma camera consists of three major parts:

- a collimator,
- a detector, and
- a set of PMTs.

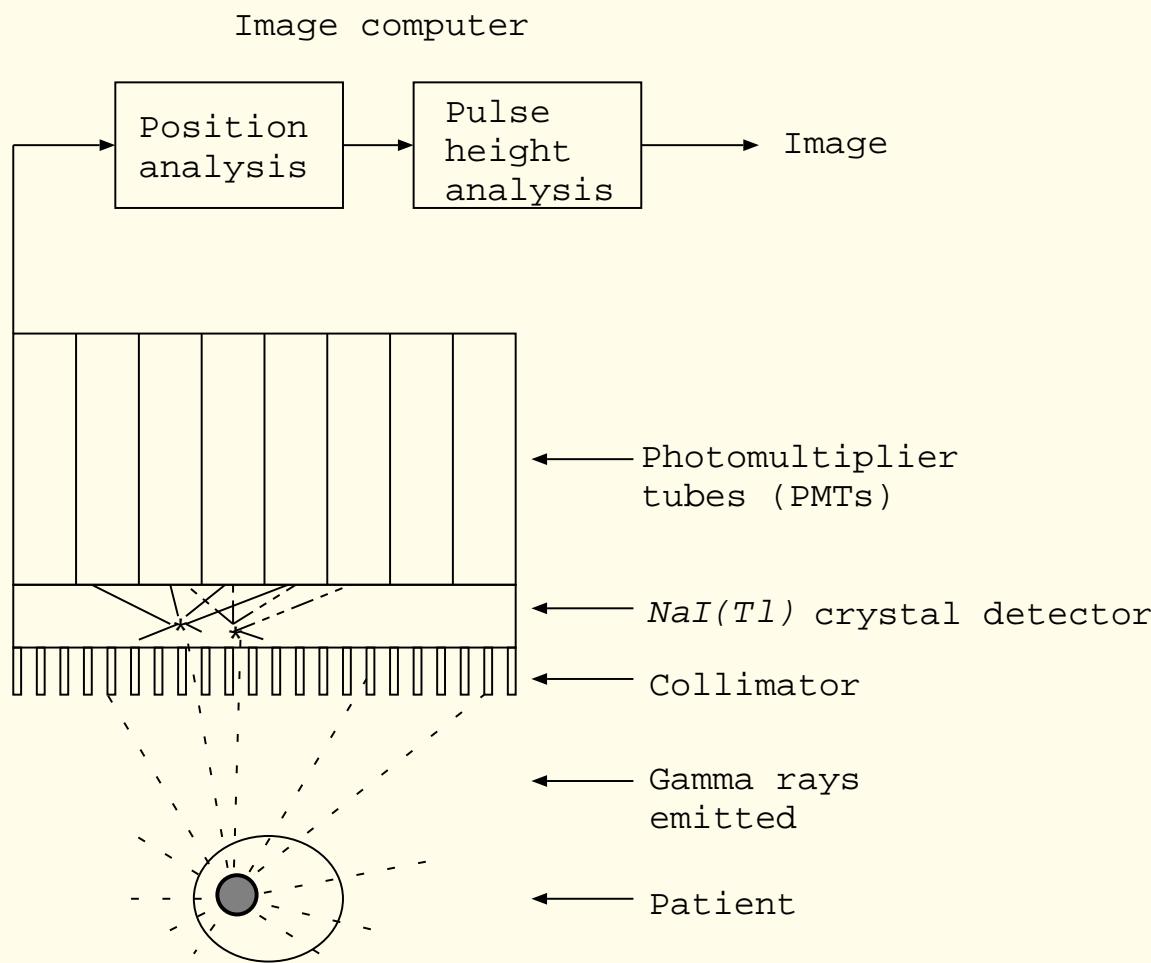


Figure 1.25: Schematic (vertical sectional) representation of a nuclear medicine imaging system with an Anger camera.



▪ Collimator:

Array of holes separated by lead septa.

Allow passage of the gamma rays that arrive along a certain path of propagation, and to block (absorb) all gamma rays outside a narrow solid angle of acceptance.

Made of lead alloys, but other materials such as tantalum, tungsten, and gold have also been used.

Different geometries of collimator holes: triangular, square, hexagonal, and round patterns.



Two key factors in collimator design:

- Geometric efficiency — the fraction of the gamma-ray photons from the source that are transmitted through the collimator to the detector.
- Geometric (spatial) resolution.

In general, for a given type of collimator, the higher the efficiency, the poorer is the resolution.

The resolution of a collimator is increased if the size of the holes is reduced or if the collimator thickness is increased.

However, these measures decrease the number of photons that will reach the crystal, and hence reduce the sensitivity and efficiency of the system.



The efficiency of a typical collimator is about 0.01%: only 1 in every 10,000 photons emitted is passed by the collimator to the crystal.

The most commonly used type of collimator is the parallel-hole collimator.

Other designs include diverging, converging, fan-beam, and pin-hole collimators.



- **Detector:**

Usually a $NaI(Tl)$ crystal of $6 - 13\text{ mm}$ thickness.

The crystal absorbs the gamma-ray photons that pass through the collimator holes, and reemits their energy as visible light (scintillation).

The thickness of the crystal determines the absorbed fraction of the gamma-ray photons by the photoelectric effect.

A thick crystal has better absorption than a thin crystal; however, a thick crystal scatters and absorbs the light before it reaches the back surface of the crystal.

A crystal of thickness 10 mm absorbs about 92% of the photons received at 140 keV .



- **Photomultiplier tubes:**

The crystal is optically coupled at its back surface to an array of PMTs.

Scintillations within the crystal are converted by the photocathodes at the front of the PMTs to photoelectrons, which are accelerated toward each of a series of dynodes held at successively higher potentials until they reach the anode at the back of the tube.

The photoelectrons produce a number of secondary electrons at each dynode, leading to a current gain of the order of 10^6 .



- **Image computer:**

The current pulses produced by the PMTs in response to scintillations in the crystal are applied to a resistor matrix that computes the points of arrival of the corresponding gamma-ray photons.

The amplitudes of the pulses represent the energy deposited by the gamma rays.

A pulse-height analyzer is used to select pulses that are within a preset energy window corresponding to the peak energy of the gamma rays.

The pulse-selection step reduces the effect of scattered rays at energy levels outside the energy window used.



The major disadvantages of SPECT are poor spatial resolution and high noise content.

The intrinsic resolution of a typical gamma camera (crystal) is $3 - 5 \text{ mm}$.

The net resolution including the effect of the collimator, expressed as the full width at half the maximum (FWHM) of the image of a thin line source (the line spread function or LSF) is $7.5 - 10 \text{ mm}$.

The main causes of noise are quantum mottle due to the low number of photons used to create images, and the random nature of gamma ray emission.

Structured noise may also be caused by nonuniformities in the gamma camera.



Single-photon emission computed tomography:

SPECT scanners usually gather 64 or 128 projections spanning 180° or 360° around the patient.

Individual scan lines from the projection images may then be processed through a reconstruction algorithm to obtain 2D sectional images.

Coronal, sagittal, and oblique sections may then be created from the 3D dataset.



Circular scanning is commonly used to acquire projection images of the body at different angles.

However, some clinical studies use elliptical scanning so as to keep the camera close to the body.

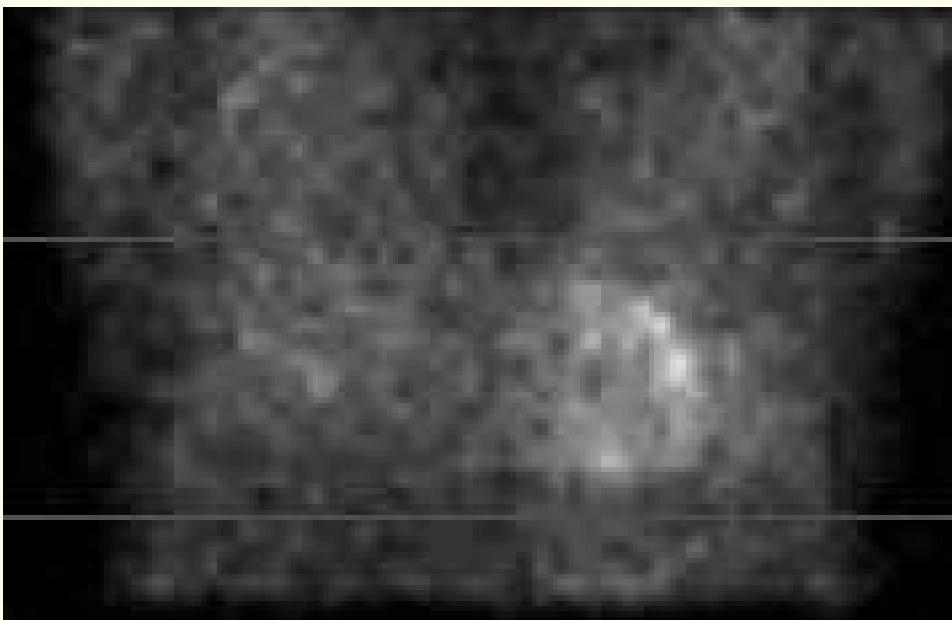
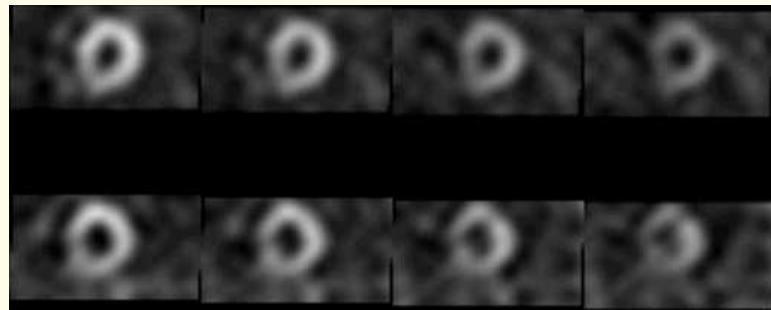
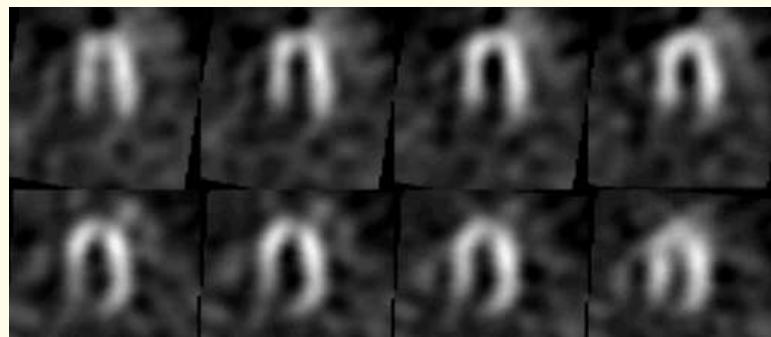


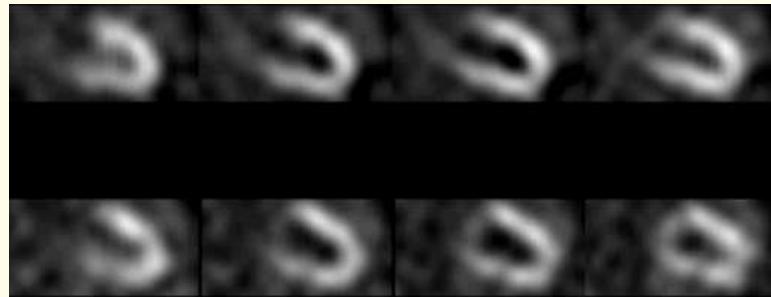
Figure 1.26: A planar or projection image of a patient used for myocardial SPECT imaging. The two horizontal lines indicate the limits of the data used to reconstruct the SPECT images shown in Figure 1.27. Image courtesy of Foothills Hospital, Calgary.



(a)



(b)



(c)

Figure 1.27: SPECT imaging of the left ventricle. (a) Short-axis images. (b) Horizontal long axis images. (c) Vertical long axis images. In each case, the upper panel shows four SPECT images after exercise (stress), and the lower panel shows the corresponding views before exercise (rest). Images courtesy of Foothills Hospital, Calgary.



Positron emission tomography (PET):

Certain isotopes of carbon (^{11}C), nitrogen (^{13}N), oxygen (^{15}O), and fluorine (^{18}F) emit positrons and are suitable for nuclear medicine imaging.

PET is based upon the simultaneous detection of the two annihilation photons produced at 511 keV and emitted in opposite directions when a positron loses its kinetic energy and combines with an electron: coincidence detection.



In one mode of data collection, a ring of bismuth-germanate detectors is used to gather emission statistics that correspond to a projection of a transversal section.

Spatial resolution: typically 5 mm .

Useful in functional imaging and physiological analysis of organs.



1.8 Ultrasonography

Ultrasound in the frequency range of $1 - 20 \text{ MHz}$ is used in diagnostic ultrasonography.

The velocity of propagation of ultrasound through a medium depends upon its compressibility:

lower compressibility results in higher velocity.

Typical velocities in human tissues:

- 330 m/s in air (the lungs);
- $1,540 \text{ m/s}$ in soft tissue; and
- $3,300 \text{ m/s}$ in bone.



A wave of ultrasound may get reflected, refracted, scattered, or absorbed as it propagates through a body.

Most modes of diagnostic ultrasonography are based upon the reflection of ultrasound at tissue interfaces.

A gel is used to minimize the presence of air between the transducer and the skin to avoid reflection at the skin surface.



Typically, pulses of ultrasound of about $1 \mu s$ duration at a repetition rate of about $1,000 \text{ pps}$ (pulses per second) are applied, and the resulting echoes are used for locating tissue interfaces and imaging.

Large, smooth surfaces in a body cause specular reflection, whereas rough surfaces and regions cause nonspecular reflection or diffuse scatter.



The normal liver, for example, is made up of clusters of parenchyma that are of the order of 2 mm in size.

Considering an ultrasound signal at 1 MHz and assuming a propagation velocity of $1,540\text{ m/s}$, the wavelength is 1.54 mm : of the order of the size of parenchymal clusters.

For this reason, ultrasound is scattered in all directions by the liver, which appears with a speckled texture.



Fluid-filled regions such as cysts have no internal structure, generate no echoes except at their boundaries, and appear as black regions on ultrasound images.

Absorption of ultrasound by bone causes shadowing in images:

tissues past bones and dense objects along the path of propagation of the beam are not imaged accurately.



The quality of ultrasonographic images is affected by

- multiple reflections,
- speckle noise due to scattering, and
- spatial distortion due to refraction.

Spatial resolution of ultrasound images: $0.5 - 3 \text{ mm}$.



- **A mode:**

A single transducer is used in this mode.

The amplitude (A) of the echoes is displayed on the vertical axis, with the corresponding depth (related to the time of arrival of the echo) on the horizontal axis.

The A mode is useful in distance measurement (ranging), with applications in the detection of retinal detachment and the detection of shift of the midline of the brain.



- **M mode:**

This mode produces a display with time on the horizontal axis and echo depth on the vertical axis.

The M mode is useful in the study of movement or motion (M), with applications in cardiac valve motion analysis.



- **B mode:**

An image of a 2D section or slice of the body is produced by using a single transducer to scan the region of interest or by using an array of sequentially activated transducers.

Real-time imaging is possible at $15 - 40 \text{ fps}$.

The B mode is useful in studying large organs, such as the liver, and in fetal imaging.



- **Doppler ultrasound:**

Based upon the change in frequency of the investigating beam caused by a moving target (the Doppler effect).

Useful in imaging blood flow.

Detection of turbulence and retrograde flow: useful in the diagnosis of stenosis or insufficiency of cardiac valves and plaques in blood vessels.

Doppler imaging may be used to obtain a combination of anatomic information with B-mode imaging and flow information obtained using pulsed Doppler.



▪ Special probes:

A variety of probes have been developed for ultrasonography of specific organs and for special applications:

- endovaginal probes for fetal imaging,
- transrectal probes for imaging the prostate,
- transesophageal probes for imaging the heart via the esophagus, and
- intravascular probes for the study of blood vessels.

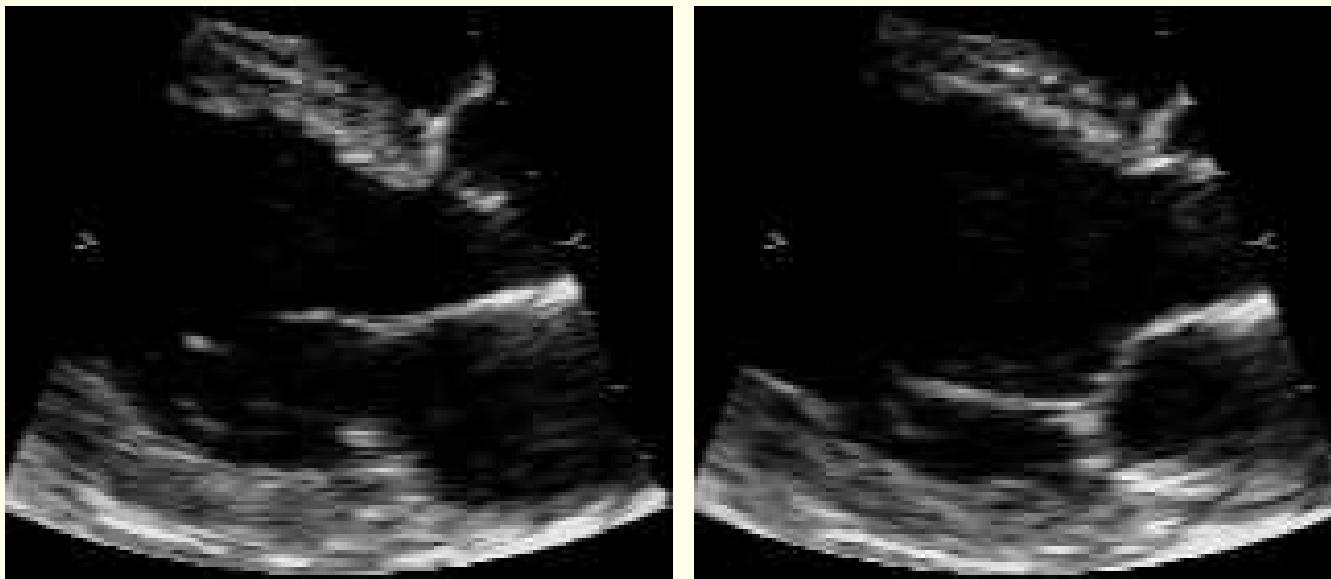


Examples:

Echocardiography — ultrasonography for the assessment of the functional integrity of heart valves.

An array of ultrasound transducers is used in the B mode to obtain a video illustrating the opening and closing activities of the valve leaflets.

Useful in the detection of stenosis and loss of flexibility of the cardiac valves due to calcification.



(a)

(b)

Figure 1.28: Two frames of the echocardiogram of a subject with normal function of the mitral valve. (a) Mitral valve in the fully open position. (b) Mitral valve in the closed position. Images courtesy of Foothills Hospital, Calgary.

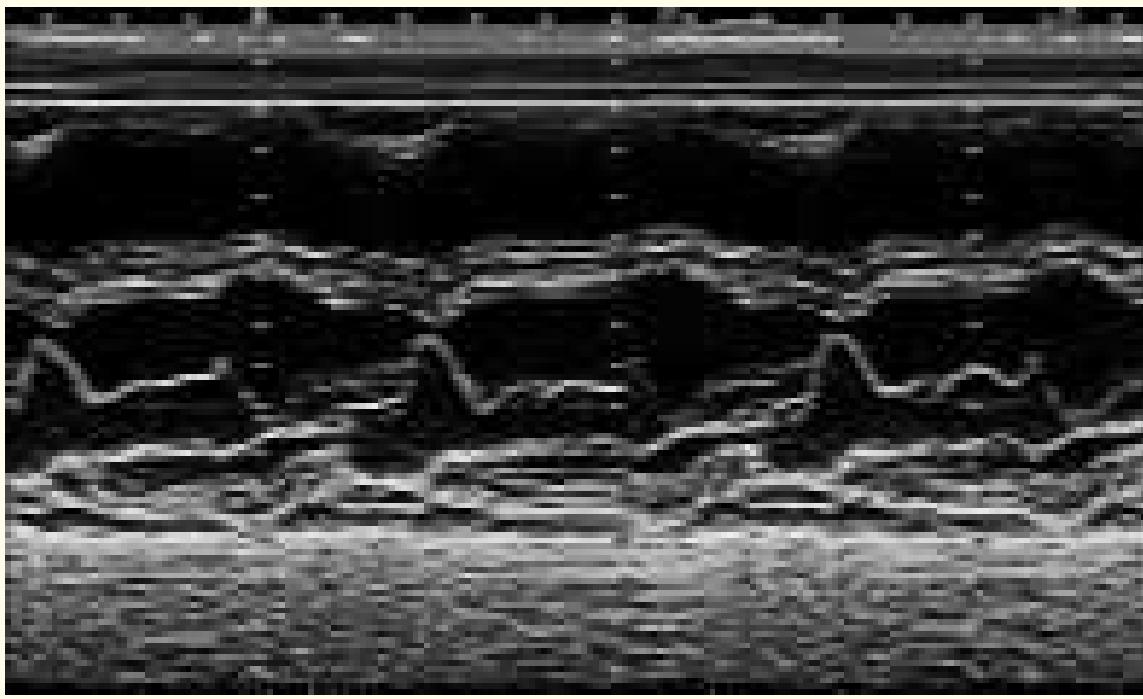


Figure 1.29: M-mode ultrasound image of a subject with normal function of the mitral valve. The horizontal axis represents time. The echo signature of the mitral valve leaflets as they open and close is illustrated. Image courtesy of Foothills Hospital, Calgary.



In spite of limitations in image quality and resolution, ultrasonography is an important medical imaging modality due to the nonionizing nature of the medium.

Ultrasonography is particularly useful in fetal imaging.

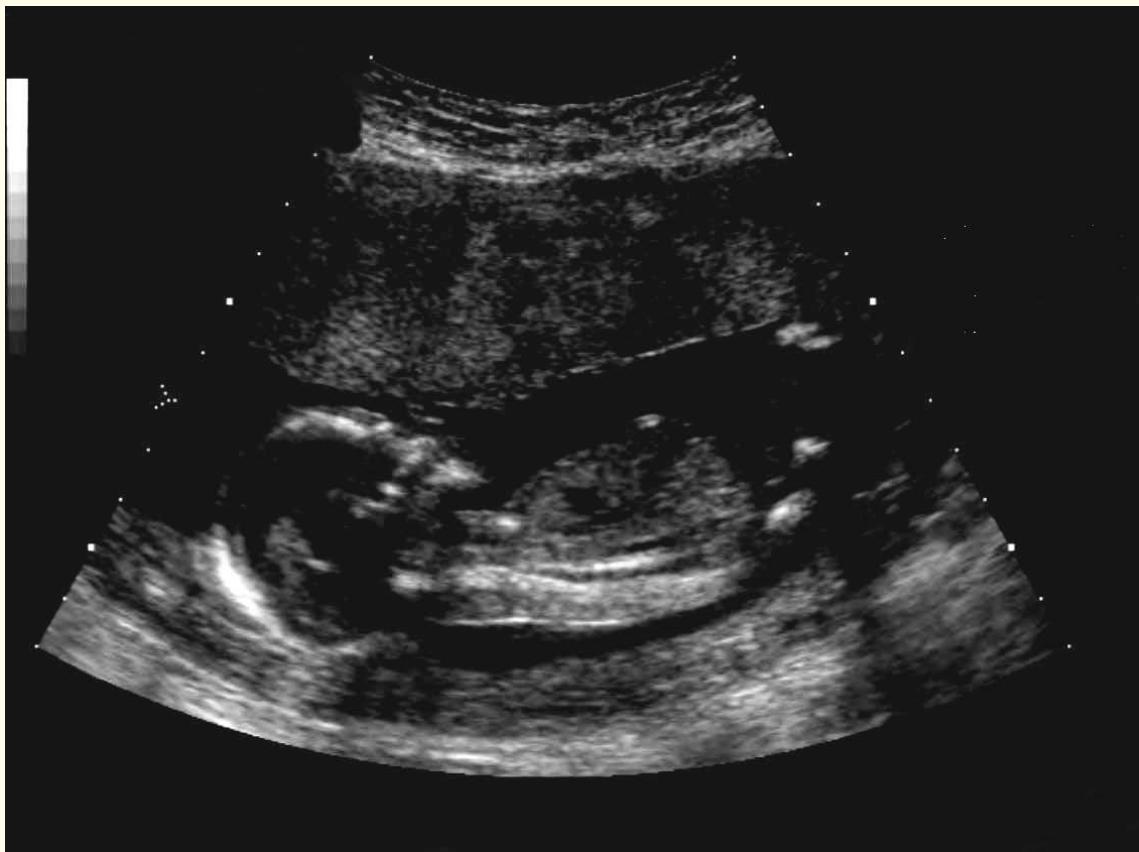


Figure 1.30: B-mode ultrasound (3.5 MHz) image of a fetus (sagittal view). Image courtesy of Foothills Hospital, Calgary.



Ultrasonography is also useful in

- tomographic imaging,
- discriminating between solid masses and fluid-filled cysts in the breast, and
- tissue characterization.



1.9 Magnetic Resonance Imaging

MRI is based on nuclear magnetic resonance (NMR):

the behavior of nuclei under the influence of externally applied magnetic and EM (RF) fields.

A nucleus with an odd number of protons or neutrons has an inherent nuclear spin and exhibits a magnetic moment: such a nucleus is said to be NMR-active.



The commonly used modes of MRI rely on:

- hydrogen (1H or proton),
- carbon (^{13}C),
- fluorine (^{19}F), or
- phosphorus (^{31}P).



In the absence of an external magnetic field, the vectors of magnetic moments of active nuclei have random orientations, resulting in no net magnetism.

When a strong external magnetic field H_o is applied (z axis), some of the nuclear spins of active nuclei align with the field (parallel or antiparallel):

the forced alignment results in a net magnetization vector.



The magnetic spin vector of each active nucleus precesses about the z axis at a frequency known as the Larmor frequency:

$$\omega_o = \gamma H_o, \quad (1.4)$$

where γ is the gyromagnetic ratio of the nucleus considered.

For protons, $\gamma = 42.57 \text{ MHz T}^{-1}$.



MRI involves controlled perturbation of the precession of nuclear spins, and measurement of the RF signals emitted when the perturbation is stopped and the nuclei return to their previous state of equilibrium.

MRI is an intrinsically 3D imaging procedure.

No mechanical scanning is involved: slice selection and scanning are performed electronically by the use of magnetic field gradients and RF pulses.



Main components and principles of MRI:

- **A magnet:** strong, uniform field, $0.5 - 4\text{ T}$.

The stronger the magnetic field, the more spins are aligned in the parallel state versus the antiparallel state, and the higher will be the signal-to-noise ratio (SNR) of the data acquired.



- An RF transmitter to deliver an RF electromagnetic pulse H_1 to the body being imaged.

The RF pulse provides a perturbation: it causes the axis of precession of the net spin vector to deviate or “flip” from the z axis.

The frequency of the RF field must be the same as that of precession of the active nuclei, such that the nuclei can absorb energy from the RF field (“resonance”).

The frequency of RF-induced rotation is given by

$$\omega_1 = \gamma H_1. \quad (1.5)$$

When the RF perturbation is removed, the active nuclei return to their unperturbed states (alignment with H_o) through various relaxation processes, emitting energy in the form of RF signals.



- A **gradient system** to apply to the body a controlled space-variant and time-variant magnetic field

$$h(t, \mathbf{x}) = \mathbf{G}(t) \cdot \mathbf{x}, \quad (1.6)$$

where \mathbf{x} is a vector representing the spatial coordinates, \mathbf{G} is the gradient applied, and t is time.

The components of \mathbf{G} along the z direction as well as in the x and y directions are controlled individually (the $x - y$ plane is orthogonal to the z axis).



The gradient causes nuclei at different positions to precess at different frequencies, and provides for spatial coding of the signal emitted from the body.

The Larmor frequency at \mathbf{x} is given by

$$\omega(\mathbf{x}) = \gamma(H_o + \mathbf{G} \cdot \mathbf{x}). \quad (1.7)$$

Nuclei at specific positions in the body may be excited selectively by applying RF pulses of specific frequencies.

The combination of the gradient fields and the RF pulses applied is called the pulse sequence.



- An RF detector system to detect the RF signals emitted from the body.

The RF signal measured outside the body represents the sum of the RF signals emitted by active nuclei from a certain part or slice of the body, as determined by the pulse sequence.

The spectral spread of the RF signal due to the application of gradients provides information on the location of the corresponding source nuclei.



- **A computing and imaging system** to reconstruct images from the measured data, as well as process and display the images.

Depending upon the pulse sequence applied, the RF signal sensed may be formulated as the 2D or 3D Fourier transform of the image to be reconstructed.

The data measured correspond to samples of the 2D Fourier transform of a sectional image located on concentric squares or circles.

The Fourier method of image reconstruction from projections may then be used to obtain the image.



Various pulse sequences may be used to measure different parameters of the tissues in the body being imaged.

The image obtained is a function of the nuclear spin density in space and the corresponding parameters of the relaxation processes involved.

Longitudinal magnetization: component of the magnetization vector along the direction of the external magnetic field.

Longitudinal relaxation: process by which longitudinal magnetization returns to its state of equilibrium (realignment with the external magnetic field) after an excitation pulse.

The time constant of longitudinal relaxation: T_1 .



A 90° RF pulse causes the net magnetization vector to be oriented in the plane perpendicular to the external magnetic field: this is known as transverse magnetization.

When the excitation is removed, the affected nuclei return to their states of equilibrium, emitting a signal, known as the free-induction decay (FID) signal, at the Larmor frequency.

Decay time constant of transverse magnetization: T_2 .

Range of T_1 for various types of tissues: 200 ms to 2,000 ms.

Range of T_2 values: 80 ms to 180 ms.



MRI is suitable for functional imaging.

The increased supply of oxygen (or oxygenated blood) to certain regions of the brain due to related stimuli may be recorded on MR images.

The difference between the prestimulus and post-stimulus images may then be used to analyze the functional aspects of specific regions of the brain.



Figure 1.31: Sagittal section of the MR image of a patient's knee. Image courtesy of Foothills Hospital, Calgary.

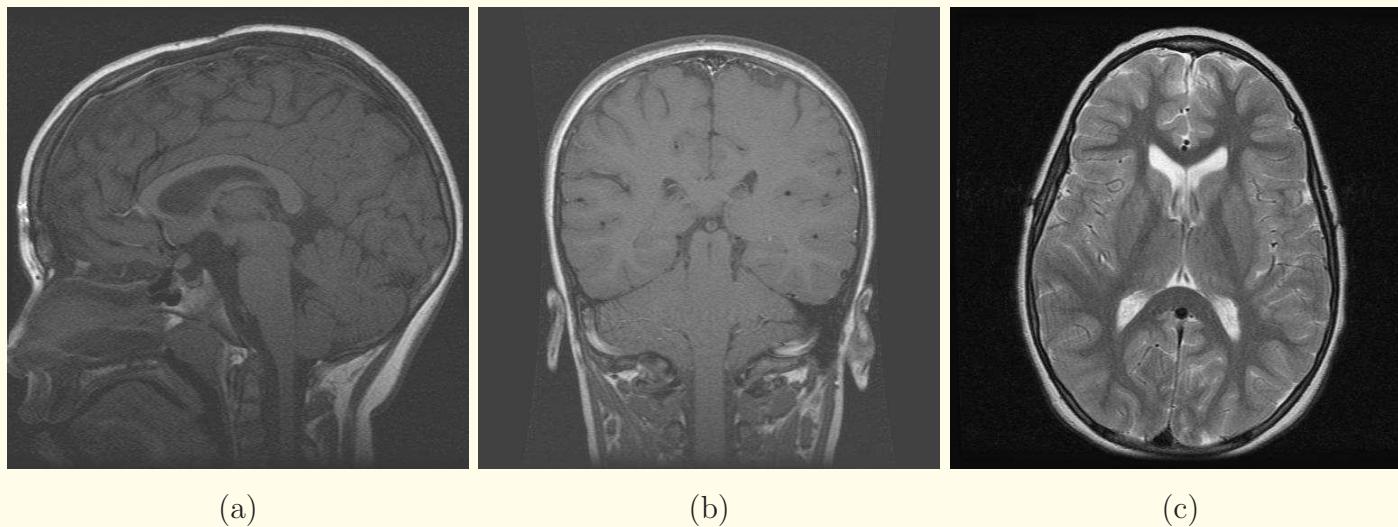


Figure 1.32: (a) Sagittal, (b) coronal, and (c) transversal (cross-sectional) MR images of a patient's head. Images courtesy of Foothills Hospital, Calgary.



1.10 Objectives of Biomedical Image Analysis

The representation of biomedical images in electronic form facilitates the following:

- Computer processing and analysis of the data.
- Computer-aided diagnosis (CAD).
- Image-guided surgery.
- Image-guided therapy.

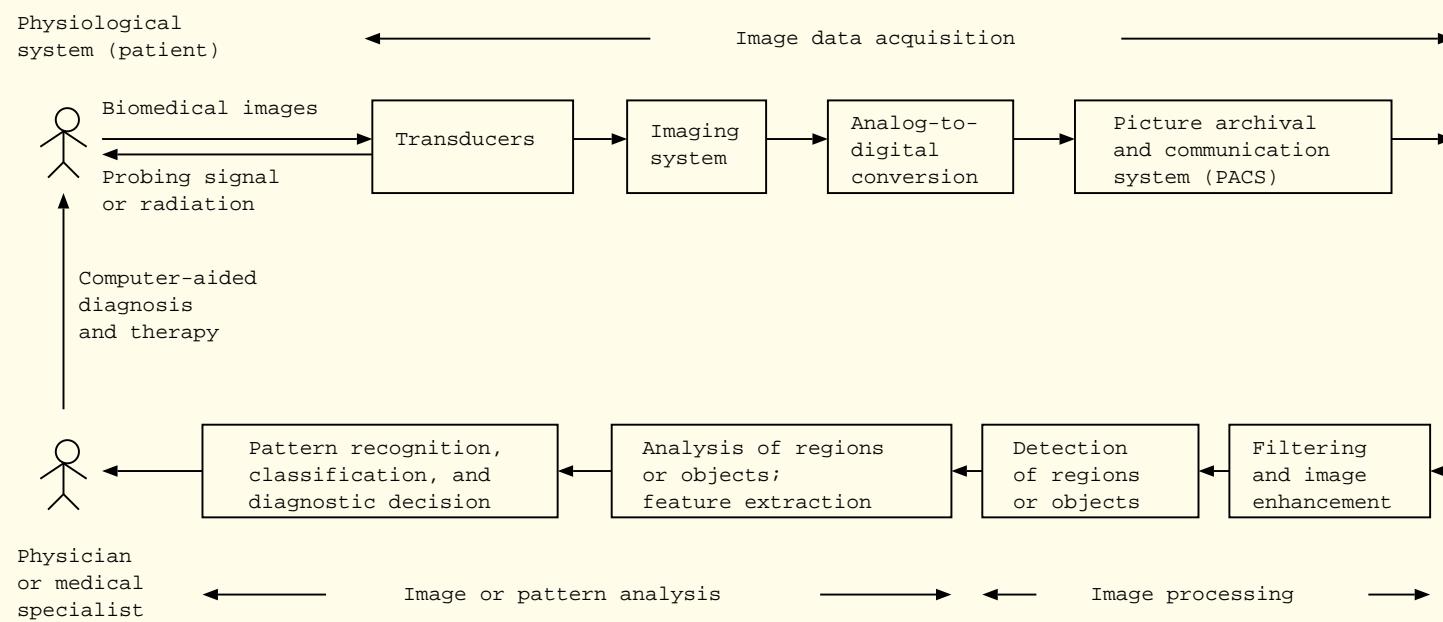


Figure 1.33: Computer-aided diagnosis and therapy based upon biomedical image analysis.



The human–instrument system:

- *The subject (or patient)*: The main purpose of biomedical imaging and image analysis is to provide a certain benefit to the subject or patient.

All systems and procedures should be designed so as not to cause undue inconvenience to the subject, and not to cause any harm or danger.

In applying invasive or risky procedures, it is extremely important to perform a risk–benefit analysis and determine if the anticipated benefits of the procedure are worth placing the subject at the risks involved.



- ***Transducers:*** films, scintillation detectors, fluorescent screens, solid-state detectors, piezoelectric crystals, X-ray generators, ultrasound generators, EM coils, electrodes, sensors.
- ***Signal-conditioning equipment:*** PMTs, amplifiers, filters.
- ***Display equipment:*** oscilloscopes, strip-chart or paper recorders, computer monitors, printers.
- ***Recording, data processing, and transmission equipment:*** films, analog-to-digital converters (ADCs), digital-to-analog converters (DACs), digital tapes, compact disks (CDs), diskettes, computers, telemetry systems, picture archival and communication systems (PACS).
- ***Control devices:*** power supply stabilizers and isolation equipment, patient intervention systems.



Major objectives of biomedical instrumentation:

- *Information gathering* — measurement of phenomena to interpret an organ, a process, or a system.
- *Screening* — investigating a large asymptomatic population for the incidence of a certain disease (early detection).
- *Diagnosis* — detection or confirmation of malfunction, pathology, or abnormality.
- *Monitoring* — obtaining periodic information about a system.
- *Therapy and control* — modification of the behavior of a system based upon the outcome of the activities listed above to ensure a specific result.
- *Evaluation* — objective analysis to determine the ability to meet functional requirements, obtain proof of performance, perform quality control, or quantify the effect of treatment.



Invasive versus noninvasive procedures:

Invasive procedures involve the placement of devices or materials inside the body, such as the insertion of

- endoscopes,
- catheter-tip sensors,
- X-ray contrast media.

Noninvasive procedures are desirable in order to minimize risk to the subject.



Active versus passive procedures:

Active data acquisition procedures require external stimuli to be applied to the subject, or require the subject to perform a certain activity to stimulate the system of interest in order to elicit the desired response.

For example, in SPECT investigations of myocardial ischemia, the patient performs vigorous exercise on a treadmill.

An ischemic zone is better delineated in SPECT images taken when the cardiac system is under stress than when at rest.



Stressing an unwell system may cause pain, irreparable damage, or death.

The investigator should be aware of such risks, perform a risk–benefit analysis, and be prepared to handle or manage adverse reactions.

Passive procedures do not require the subject to perform any activity.



Most organizations require ethical approval by specialized committees for experimental procedures involving human or animal subjects, with the aim of

- minimizing the risk and discomfort to the subject, and
- maximizing the benefits to both the subject and the investigator.



1.11 Computer-aided Diagnosis (CAD)

Radiologists, physicians, cardiologists, neuroscientists, pathologists, and other health-care professionals are highly trained and skilled practitioners.

Why then would we want to suggest the use of computers for the analysis of biomedical images?



- Humans are highly skilled and fast in the analysis of visual patterns, but are slow (usually) in arithmetic operations with large numbers of values.

Computers can perform millions of arithmetic operations or computations per second.

However, the recognition of objects and patterns in images using mathematical procedures requires huge numbers of operations: slow response from low-level computers.

A trained human observer, can usually recognize an object or a pattern in an instant.



- Humans could be affected by fatigue, boredom, and environmental factors: susceptible to committing errors.
Working with large numbers of images in one sitting, such as in breast cancer screening, poses practical difficulties.
A human observer could be distracted by other events in the surrounding areas and may miss uncommon signs present in some images.
- Computers, being inanimate but mathematically accurate and consistent machines, can be designed to perform computationally specific and repetitive tasks.



- Analysis by humans is usually subjective and qualitative.
Computers can assist in quantitative and objective analysis.
Quantitative analysis becomes possible by the application of computers to biomedical images.
The logic of medical or clinical diagnosis via image analysis could then be *objectively* encoded and *consistently* applied in routine or repetitive tasks.



- Analysis by humans is subject to interobserver as well as intraobserver variations (with time).
The former could be due to lack of diligence or due to inconsistent application of knowledge;
the latter due to variations in training and the level of understanding or competence.
Computers can apply a given procedure repeatedly and whenever recalled in a consistent manner.
It is possible to encode the knowledge (the logical processes) of many experts into a single computational procedure:
enable a computer with the collective “intelligence” of several human experts in an area of interest.



Computer-*aided* diagnosis or automated diagnosis?

A physician or medical specialist typically uses other information in addition to images, including the general physical appearance and mental state of the patient, family history, and socio-economic factors affecting the patient, many of which are not amenable to quantification and logical rule-based processes.

Biomedical images are, at best, indirect indicators of the state of the patient; many cases may lack a direct or unique image-to-pathology relationship.

The results of image analysis need to be integrated with other clinical signs, symptoms, and information by a specialist.



Above all, the *intuition* of the medical specialist plays an important role in arriving at the final diagnosis.

For these reasons, and keeping in mind the realms of practice of various licensed and regulated professions, liability, and legal factors, the final diagnostic decision is best left to the physician or medical specialist.

It could be expected that quantitative and objective analysis facilitated by the application of computers to biomedical image analysis will lead to a more accurate diagnostic decision *by the physician.*



On the importance of quantitative analysis:

“When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of *science*. ”

Lord Kelvin (William Thomson, 1824 – 1907)



On assumptions made in quantitative analysis:

“Things do not in general run around with their measure stamped on them like the capacity of a freight car; it requires a certain amount of investigation to discover what their measures are ... What most experimenters take for granted before they begin their experiments is infinitely more interesting than any results to which their experiments lead.”

Norbert Wiener (1894 – 1964)







2

Image Quality and Information Content

Several factors affect image quality and information content.

Good understanding of such factors and appropriate characterization of the concomitant loss in image quality essential in order to design image processing techniques to remove the degradation and/or improve image quality.



Inherent problem in characterizing image quality:
judged by human observers in a *subjective* manner.

To quantify the notion of image quality is a difficult proposition.

Multifaceted characteristics of information in terms of:

- statistical,
- structural,
- perceptual,
- semantic, and
- diagnostic connotations.



2.1 Difficulties in Image Acquisition and Analysis

Accessibility of the organ of interest:

Several organs of interest in imaging-based investigation are situated well within the body, encased in protective and difficult-to-access regions.

Brain: protected by the skull.

Prostate: at the base of the bladder near the pelvic outlet.

Visualization of the arteries in the brain requires the injection of an X-ray contrast agent and the subtraction of a reference image.



Special transrectal probes have been designed for 3D ultrasonic imaging of the prostate.

Images obtained as above tend to be affected by severe artifacts.



Variability of information: Biological systems exhibit great ranges of inherent variability within their different categories.

The intrinsic and natural variability presented by biological entities within a given class far exceeds the variability that we may observe in engineering, physical, and manufactured samples.

The distinction between a normal pattern and an abnormal pattern is often clouded by overlap between the ranges of the features or variables used to characterize the two categories;

the problem is compounded when multiple abnormalities need to be considered.

Imaging conditions and parameters could cause ambiguities due to the effects of subject positioning and projection.



Most malignant breast tumors are irregular and spiculated in shape, whereas benign masses are smooth and round/oval.

However, some malignant tumors may present smooth shapes, and some benign masses may have rough shapes.

A tumor may present a rough appearance in one view or projection, but a smoother profile in another.

The notion of shape roughness is nonspecific and open-ended.



Overlapping patterns caused by ligaments, ducts, and breast tissue in other planes could also affect the appearance of tumors and masses in projection images.

The use of multiple views and spot magnification imaging could help resolve some of these ambiguities, but at the cost of additional radiation dose to the subject.



Physiological artifacts and interference:

Physiological systems are dynamic and active.

Some activities, such as breathing, may be suspended voluntarily by an adult for brief periods of time to permit improved imaging.

However, cardiac activity, blood circulation, and peristaltic movement are not under one's volitional control.

An analyst should pay attention to potential physiological artifacts when interpreting biomedical images.



Energy limitations:

In X-ray mammography, considering the fact that the organ imaged is mainly composed of soft tissues, a low kVp would be desired in order to maximize image contrast.

However, low-energy X-ray photons are absorbed more readily than high-energy photons by the skin and breast tissues, thereby increasing the radiation dose to the patient.

A compromise is required between these two considerations.



Patient safety:

The protection of the subject in a study from electrical shock, radiation hazard, and other potentially dangerous conditions is an unquestionable requirement of paramount importance.

Most organizations require ethical approval by specialized committees for experimental procedures involving human or animal subjects, with the aim of

- minimizing the risk and discomfort to the subject, and
- maximizing the benefits to both the subjects and the investigator.



2.2 Characterization of Image Quality

Images are complex sources of several items of information.

Many measures available to represent quantitatively several attributes of images related to impressions of quality.



Changes in measures related to quality may be analyzed for:

- comparison of images generated by different imaging systems;
- comparison of images obtained using different imaging parameter settings of a given system;
- comparison of the results of image enhancement algorithms;
- assessment of the effect of the passage of an image through a transmission channel or medium; and
- assessment of images compressed by different data compression techniques at different rates of loss.



2.3 Digitization of Images

The representation of natural scenes and objects as digital images for processing using computers requires two steps:

- sampling, and
- quantization.

Both of these steps could potentially cause loss of quality and introduce artifacts.



2.3.1 Sampling

Sampling is the process of representing a continuous-time or continuous-space signal on a discrete grid, with samples that are separated by (usually) uniform intervals.

A band-limited signal with the frequency of its fastest component being $f_m \text{ Hz}$ may be represented without loss by its samples obtained at the Nyquist rate of $f_s = 2 f_m \text{ Hz}$.

Sampling may be modeled as the multiplication of the given analog signal with a periodic train of impulses.

The multiplication of two signals in the time domain corresponds to the convolution of their Fourier spectra.



The Fourier transform of a periodic train of impulses is another periodic train of impulses with a period that is equal to the inverse of the period in the time domain (that is, $f_s \text{ Hz}$).

Therefore, the Fourier spectrum of the sampled signal is periodic, with a period equal to $f_s \text{ Hz}$.

A sampled signal has infinite bandwidth; however, the sampled signal contains distinct or unique frequency components only up to $f_m = \pm f_s/2 \text{ Hz}$.

If the signal as above is sampled at a rate lower than $f_s \text{ Hz}$, an error known as *aliasing* occurs, where the frequency components above $f_s/2 \text{ Hz}$ appear at lower frequencies.

It then becomes impossible to recover the original signal from its sampled version.



If sampled at a rate of at least $f_s = 2 f_m \text{ Hz}$, the original signal may be recovered from its sampled version by lowpass filtering and extracting the base-band component over the band $\pm f_m \text{ Hz}$.

If an ideal (rectangular) lowpass filter were to be used, the equivalent operation in the time domain would be convolution with a sinc function (which is of infinite duration).

This operation is known as *interpolation*.



In 1D DSP, in order to prevent aliasing errors, it is common to use an *anti-aliasing filter* prior to the sampling of 1D signals, with a pass-band that is close to $f_s/2 \text{ Hz}$.

This requires the prior knowledge that the signal contains no significant energy or information beyond $f_m \leq f_s/2 \text{ Hz}$.

However, in most real-life applications of imaging and image processing, it is not possible to estimate the frequency content of the images, and also not possible to apply anti-aliasing filters.



Figure 2.1 illustrates the loss of quality associated with sampling an image at lower and lower numbers of pixels.

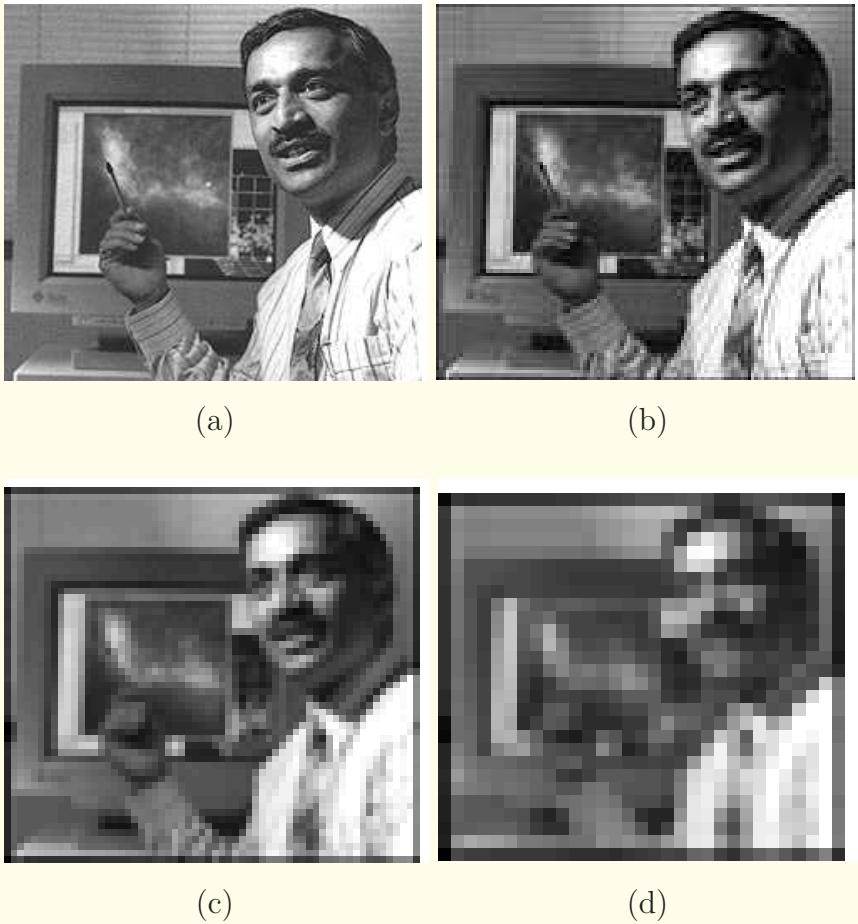


Figure 2.1: Effect of sampling on the appearance and quality of an image: (a) 225×250 pixels; (b) 112×125 pixels; (c) 56×62 pixels; and (d) 28×31 pixels. All four images have 256 gray levels at 8 bits per pixel.



2.3.2 Quantization

Quantization is the process of representing the values of a sampled signal or image using a finite set of allowed values.

Using n bits per sample and positive integers only, there exist 2^n possible *quantized* levels, spanning the range $[0, 2^n - 1]$.

If $n = 8$ bits are used to represent each pixel, there can exist 256 values or *gray levels* in the range $[0, 255]$.



It is necessary to map appropriately the range of variation of the given analog signal to the input dynamic range of the quantizer.

The decision levels of the quantizer should be optimized in accordance with the probability density function (PDF) of the original signal or image.



Lloyd–Max quantization:

Let $p(r)$ represent the PDF of the amplitude or gray levels in the given image, with the values of the continuous or analog variable r varying within the range $[r_{\min}, r_{\max}]$.

Let the range $[r_{\min}, r_{\max}]$ be divided into L parts demarcated by the decision levels $R_0, R_1, R_2, \dots, R_L$, with $R_0 = r_{\min}$ and $R_L = r_{\max}$; see Figure 2.2.

Let the L output levels of the quantizer represent the values $Q_0, Q_1, Q_2, \dots, Q_{L-1}$.



The mean-squared error (MSE) in representing the analog signal by its quantized values is given by

$$\overline{\varepsilon^2} = \sum_{l=0}^{L-1} \int_{R_l}^{R_{l+1}} (r - Q_l)^2 p(r) dr. \quad (2.1)$$

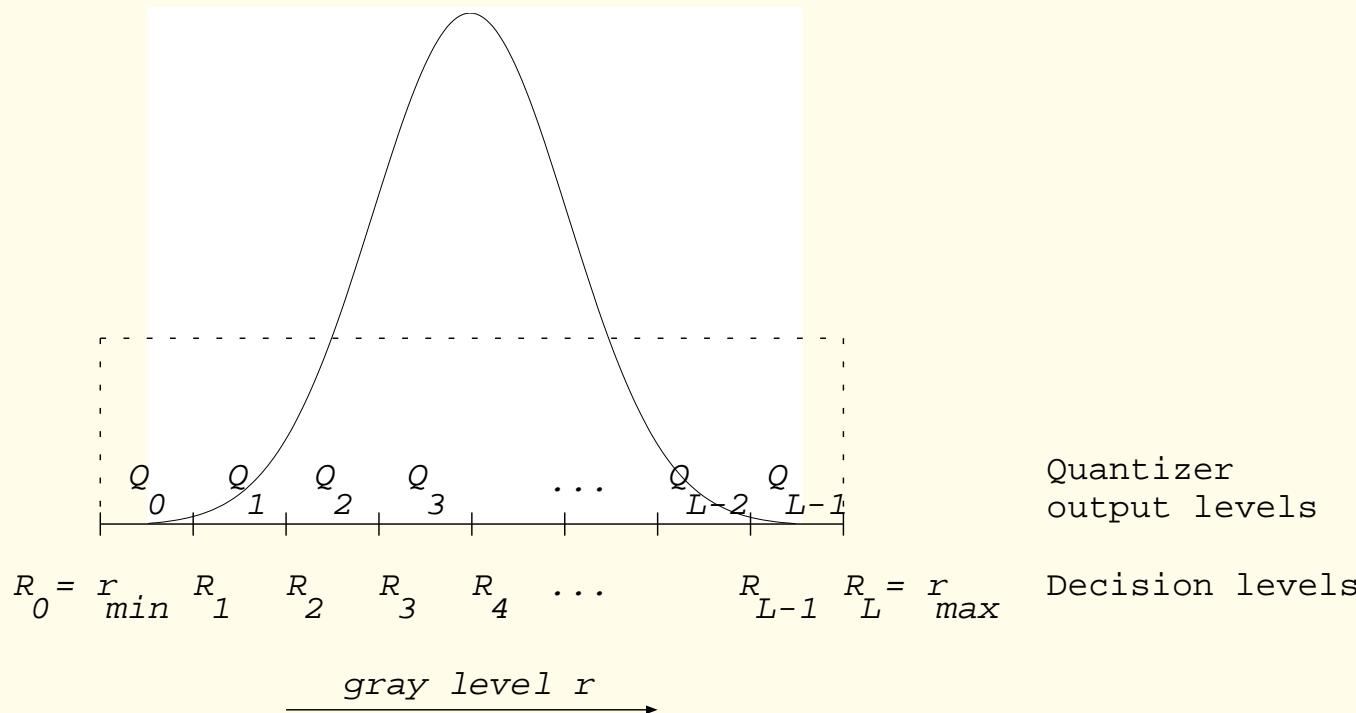


Figure 2.2: Quantization of an image gray-level signal r with a Gaussian (solid line) or uniform (dashed line) PDF. The quantizer output levels are indicated by Q_l and the decision levels represented by R_l .



A classical result indicates that the output level Q_l should lie at the centroid of the part of the PDF between the decision levels R_l and R_{l+1} , given by

$$Q_l = \frac{\int_{R_l}^{R_{l+1}} r p(r) dr}{\int_{R_l}^{R_{l+1}} p(r) dr}, \quad (2.2)$$

which reduces to

$$Q_l = \frac{R_l + R_{l+1}}{2} \quad (2.3)$$

if the PDF is uniform.

The decision levels are then given by

$$R_l = \frac{Q_{l-1} + Q_l}{2}. \quad (2.4)$$

The use of an inadequate number of quantized gray levels leads to false contours and poor representation of image intensities.



Table 2.1: Relationships Between Tissue Type, Tissue Density, X-ray Attenuation Coefficient, Hounsfield Units (*HU*), Optical Density (*OD*), and Gray Level. The X-ray Attenuation Coefficient was Measured at a Photon Energy of 103.2 keV .

Tissue type	Density gm/cm^3	X-ray atten. (cm^{-1})	Hounsfield units	Optical density	Gray level (brightness)	Appearance in image
lung	< 0.001	lower	low	high	low	dark
			[$-700, -800$]			
liver	1.2	0.18	medium	medium	medium	gray
			[$50, 70$]			
bone	1.9	higher	high	low	high	white
			[$+800, +1,000$]			

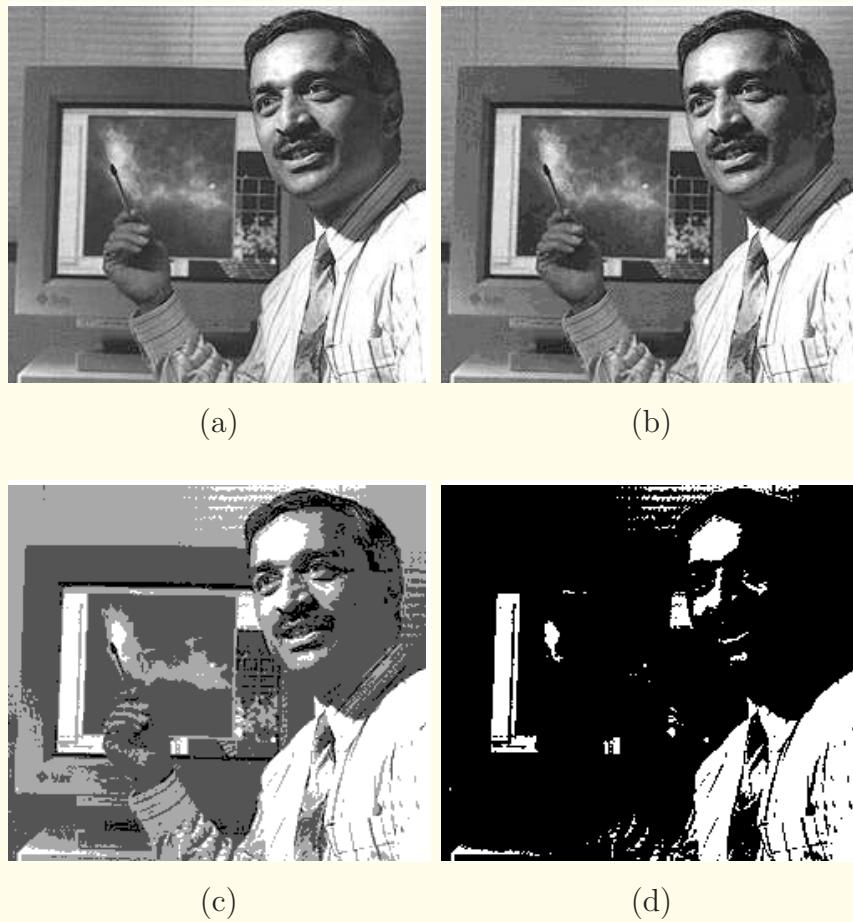


Figure 2.3: Effect of gray-level quantization on the appearance and quality of an image: (a) 64 gray levels (6 bits per pixel); (b) 16 gray levels (4 bits per pixel); (c) four gray levels (2 bits per pixel); and (d) two gray levels (1 bit per pixel) All four images have 225×250 pixels. Compare with the image in Figure 2.1 (a) with 256 gray levels at 8 bits per pixel.



2.3.3 Array and matrix representation of images

Images commonly represented as 2D functions of space: $f(x, y)$.

A digital image $f(m, n)$ may be interpreted as a discretized version of $f(x, y)$ in a 2D array, or as a matrix.

Notational differences between the representation of an image as a function of space and as a matrix: *source of confusion!*

An $M \times N$ matrix has M rows and N columns;

its height is M and width is N ;

numbering of the elements starts with $(1, 1)$ at the top-left corner and ends with (M, N) at the lower-right corner of the image.



A function of space $f(x, y)$ that has been converted into a digital representation $f(m, n)$ is typically placed in the first quadrant in the Cartesian coordinate system.

Then, an $M \times N$ will have a width of M and height of N ;

indexing of the elements starts with $(0, 0)$ at the origin at the bottom-left corner and ends with $(M - 1, N - 1)$ at the upper-right corner of the image.

The size of a matrix is expressed as *rows* \times *columns*,

the size of an image is usually expressed as *width* \times *height*.

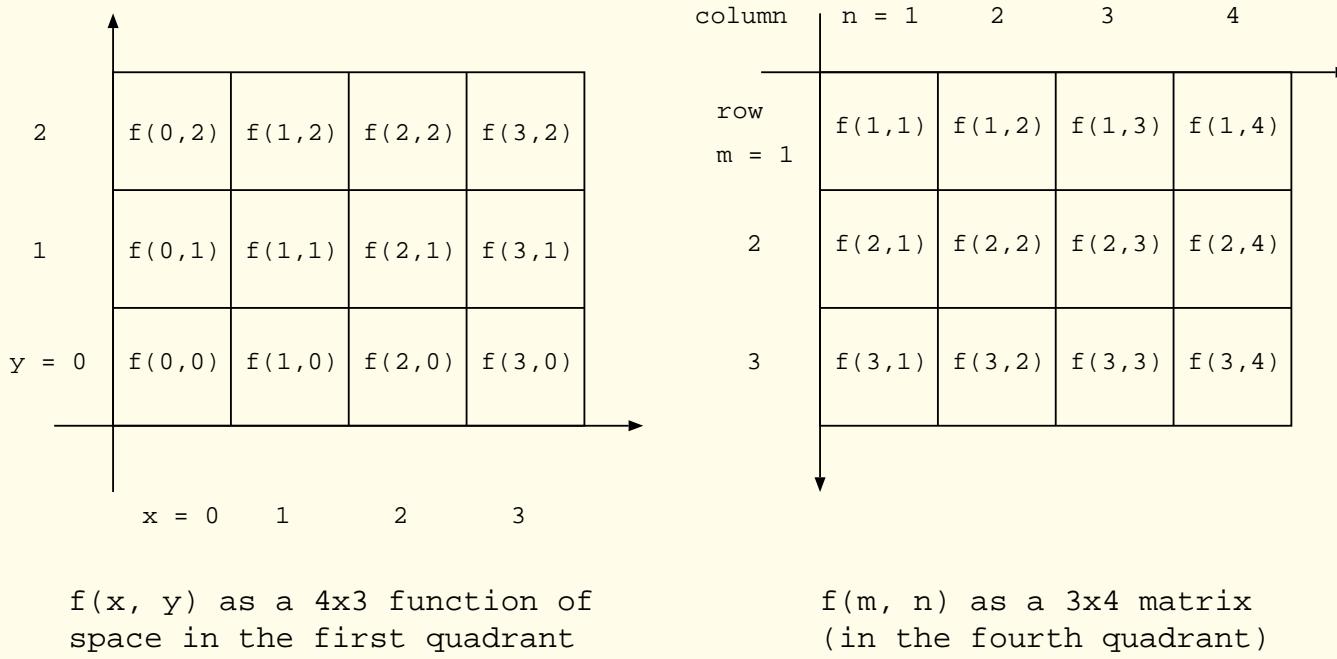


Figure 2.4: Array and matrix representation of an image.



2.4 Optical Density

The value of a picture element or cell — commonly known as a *pixel*, or occasionally as a pel — may be expressed in terms of

a physical attribute such as temperature, density, or X-ray attenuation coefficient;

the intensity of light reflected from the body at the location corresponding to the pixel;

or the transmittance at the corresponding location on a film rendition of the image.



The OD at a spot on a film is defined as

$$OD = \log_{10} \left[\frac{I_i}{I_o} \right]. \quad (2.5)$$

A perfectly clear spot will transmit all of the light that is input and will have $OD = 0$;

a dark spot that reduces the intensity of the input light by a factor of 1,000 will have $OD = 3$.

X-ray films: $OD \approx 0$ to $OD \approx 3.5$.

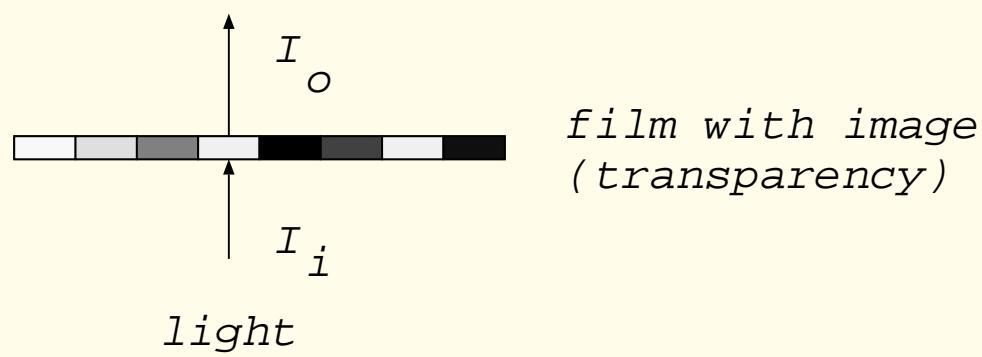


Figure 2.5: Measurement of the optical density at a spot on a film or transparency using a laser microdensitometer.



2.5 Dynamic Range

The dynamic range of an imaging system or a variable is its range or gamut of operation,

usually limited to the portion of linear response,

expressed as the maximum — minimum value of the variable.

X-ray films for mammography: dynamic range of $0 - 3.5 \text{ OD}$.

Modern CRT monitors provide dynamic range of the order of $0 - 600 \text{ cd/m}^2$ in luminance or $1 : 1,000$ in sampled gray levels.



Device A has a larger slope or “gamma” than Device B, and hence can provide higher contrast.

Device B has a larger latitude, or breadth of exposure and optical density over which it can operate, than Device A.

Plots of film density versus the log of (X-ray) exposure are known as Hurter–Driffield or H-D curves.



Air in the lungs and bowels, as well as fat in various organs including the breast, tend to extend the dynamic range of images toward the lower end of the density scale.

Bone, calcifications in the breast and in tumors, as well as metallic implants such as screws in bones and surgical clips contribute to high-density areas in images.

Mammograms: dynamic range of $0 - 3.5\text{ OD}$.

CT images: dynamic range of $-1,000$ to $+1,000\text{ HU}$.

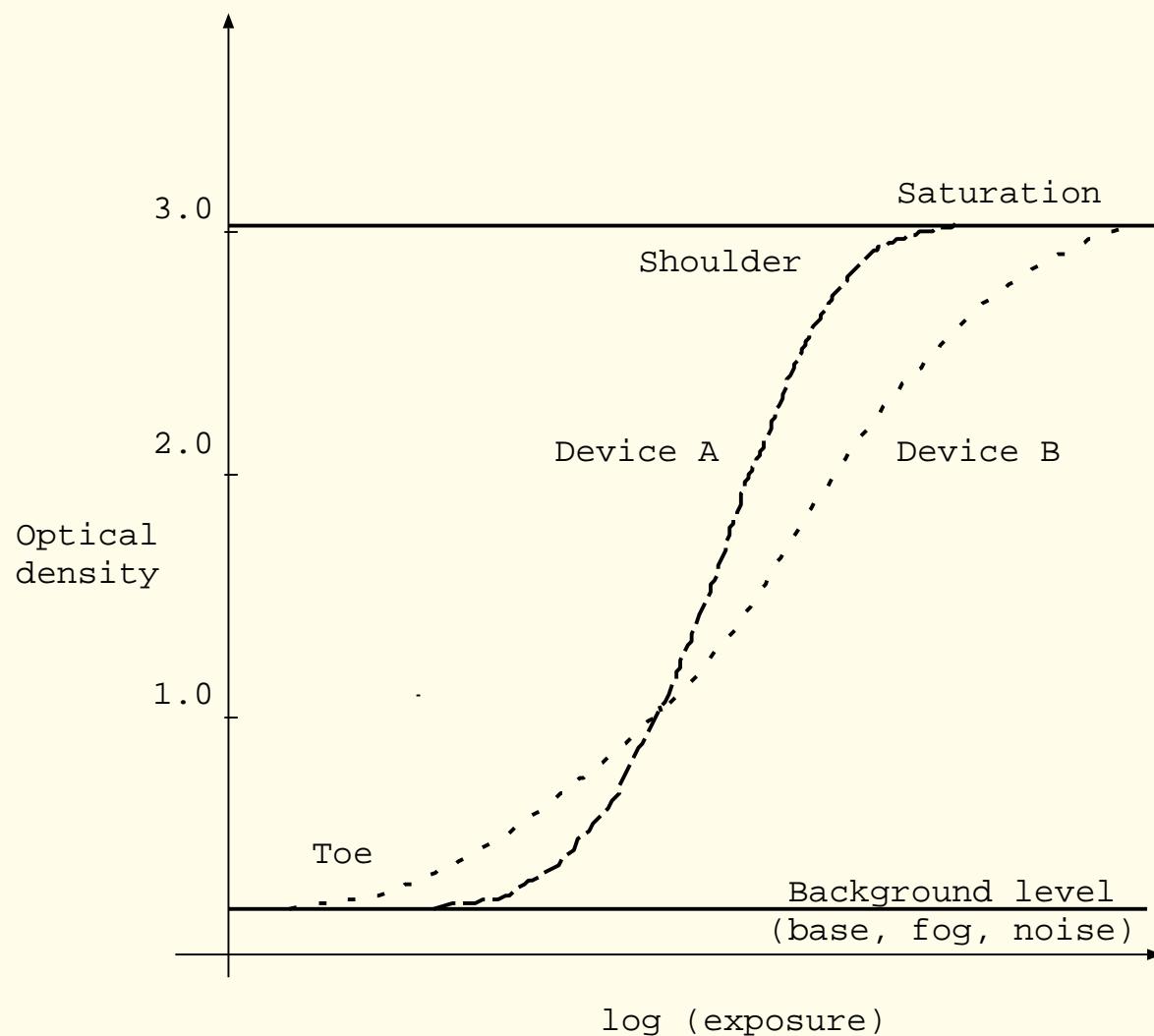


Figure 2.6: Characteristic response curves of two hypothetical imaging devices.



2.6 Contrast

$$C_{OD} = f_{OD} - b_{OD}, \quad (2.6)$$

where f_{OD} and b_{OD} represent the foreground ROI and background OD , respectively.

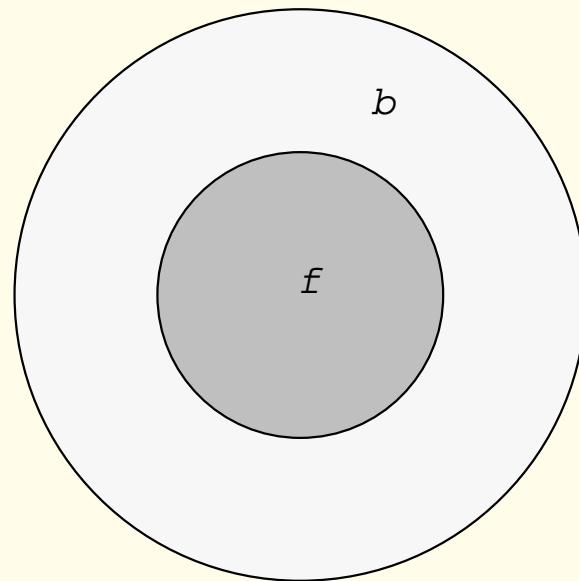


Figure 2.7: Illustration of the notion of contrast, comparing a foreground region f with its background b .



When the image parameter has not been normalized, the measure of contrast will require normalization.

If, for example, f and b represent the average light intensities emitted or reflected from the foreground ROI and the background, respectively, contrast may be defined as

$$C = \frac{f - b}{f + b}, \quad (2.7)$$

or as

$$C_1 = \frac{f - b}{b}. \quad (2.8)$$

Due to the use of a reference background, the measures defined above are often referred to as *simultaneous contrast*.



Example:

$$C_l = \frac{130 - 150}{150} = -0.1333, \quad (2.9)$$

$$C_r = \frac{130 - 50}{50} = +1.6. \quad (2.10)$$

C_l and C_r using Equation 2.7: -0.0714 and $+0.444$;

advantage: values limited to $[-1, 1]$.

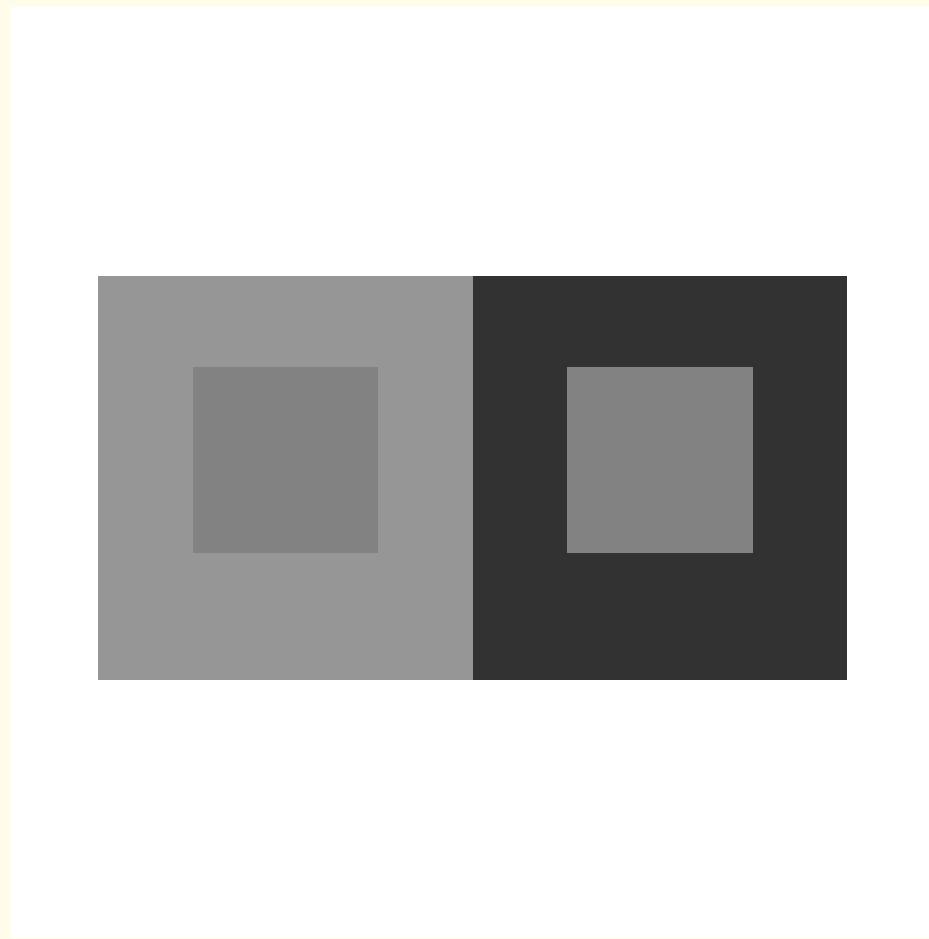


Figure 2.8: Illustration of the effect of the background on the perception of an object (simultaneous contrast). The two inner squares have the same gray level of 130, but are placed on different background levels of 150 on the left and 50 on the right.



Just-noticeable difference:

$(f - b)/b$ at the level of minimal perception of the object f for the background b .

Weber's law:

JND is almost constant ≈ 0.02 or 2% over a wide range of background intensity.



Example: The five bars have intensity values of 155, 175, 195, 215, and 235. Background: 150.

Contrast of the first bar

$$C_l = \frac{155 - 150}{150} = +0.033. \quad (2.11)$$



Figure 2.9: Illustration of the notion of just-noticeable difference. The five bars have intensity values of (from left to right) 155, 175, 195, 215, and 235, and are placed on a background of 150. The first bar is barely noticeable; the contrast of the bars increases from left to right.



Example: A calcification that appears against fat and low-density tissue may possess high contrast and be easily visible.

A similar calcification against a background of high-density breast tissue, or a calcification within a high-density tumor, could possess low contrast, and be difficult to detect.

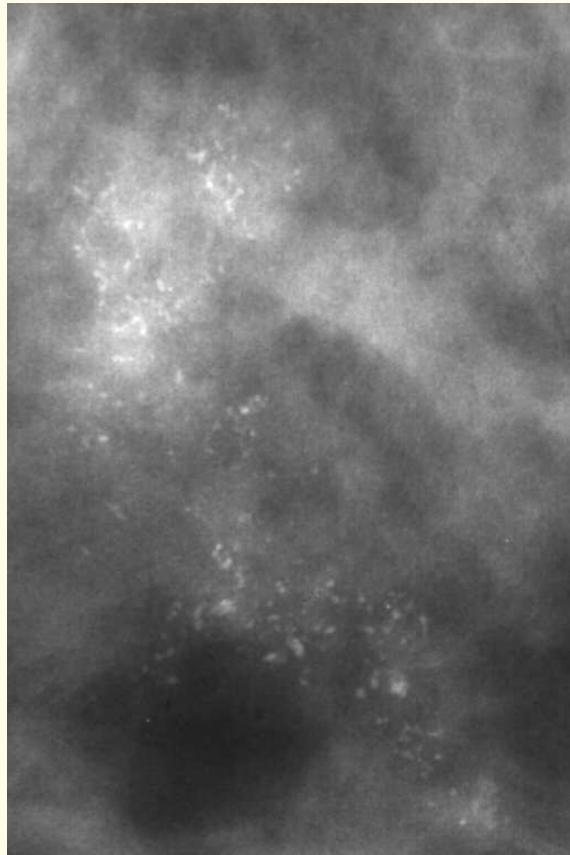


Figure 2.10: Part of a mammogram with several calcifications associated with malignant breast disease. The density of the background affects the contrast and visibility of the calcifications. The image has 768×512 pixels at a resolution of $62 \mu\text{m}$; the true width of the image is about 32 mm .



2.7 Histogram

Dynamic range: global information on the extent or spread of intensity levels across the image.

Histogram: information on the spread of gray levels over the complete dynamic range of the image across all pixels.



Consider an image $f(m, n)$ of size $M \times N$ pixels, with gray levels $l = 0, 1, 2, \dots, L - 1$.

The histogram of the image may be defined as

$$P_f(l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \delta_d[f(m, n) - l], \quad l = 0, 1, 2, \dots, L - 1, \quad (2.12)$$

where the discrete unit impulse function or delta function is

$$\delta_d(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (2.13)$$



$$\sum_{l=0}^{L-1} P_f(l) = MN. \quad (2.14)$$

The area under the function $P_f(l)$, when multiplied with an appropriate scaling factor, provides the total intensity, density, or brightness of the image,

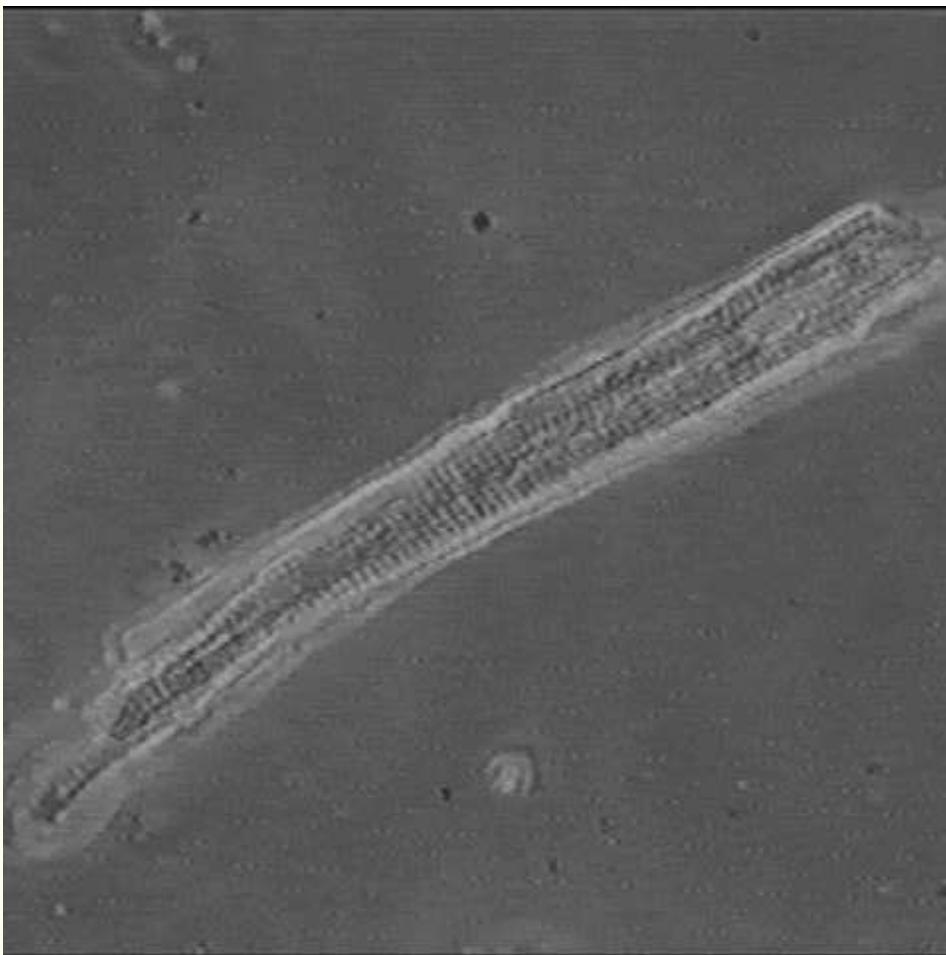
depending upon the physical parameter represented by the pixels.



The normalized histogram may be taken to represent the probability density function (PDF) $p_f(l)$ of the image-generating process:

$$p_f(l) = \frac{1}{MN} P_f(l). \quad (2.15)$$

$$\sum_{l=0}^{L-1} p_f(l) = 1. \quad (2.16)$$



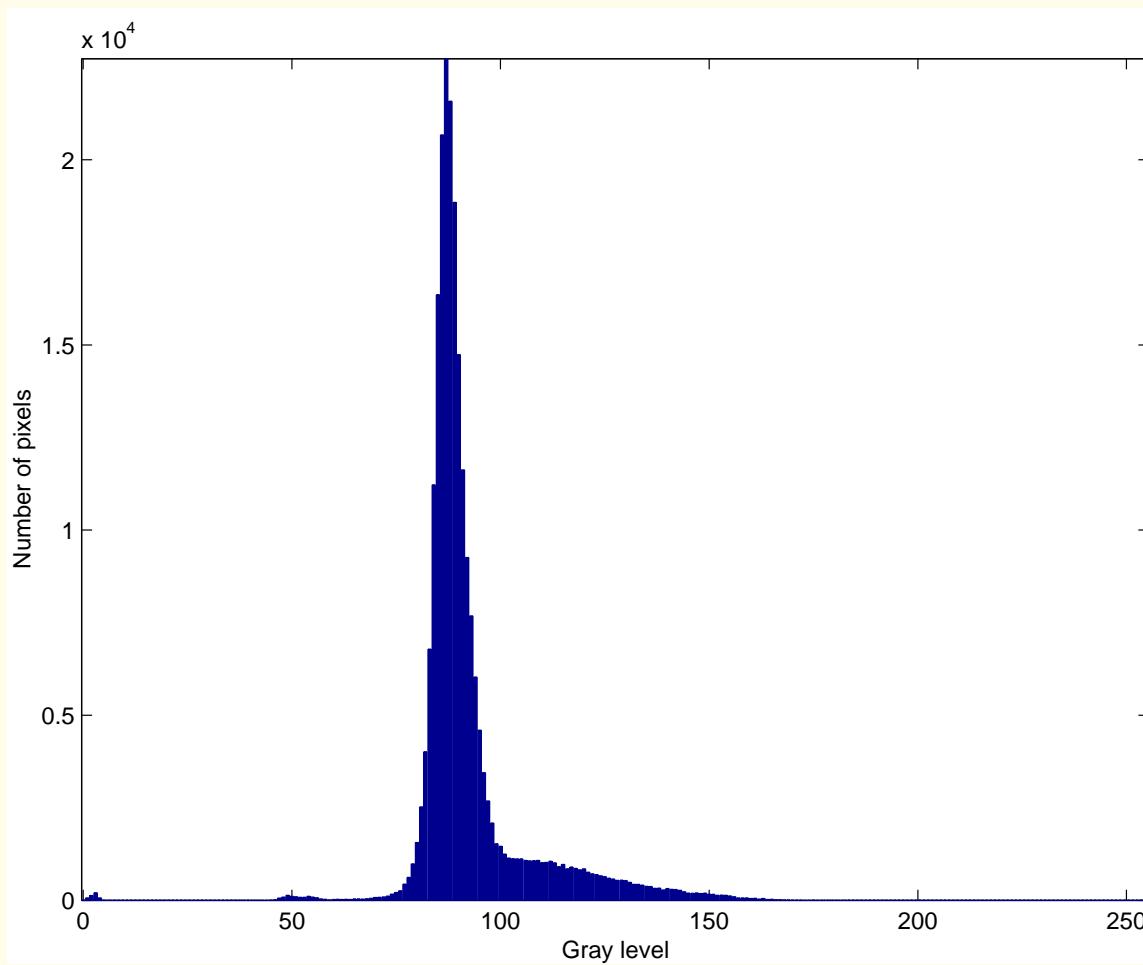
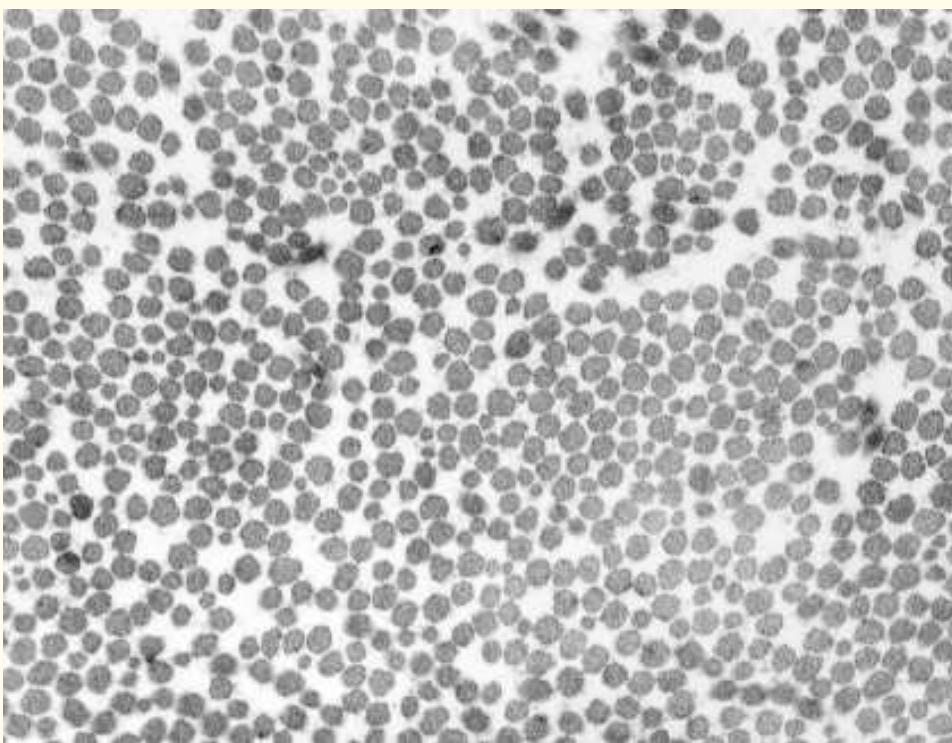
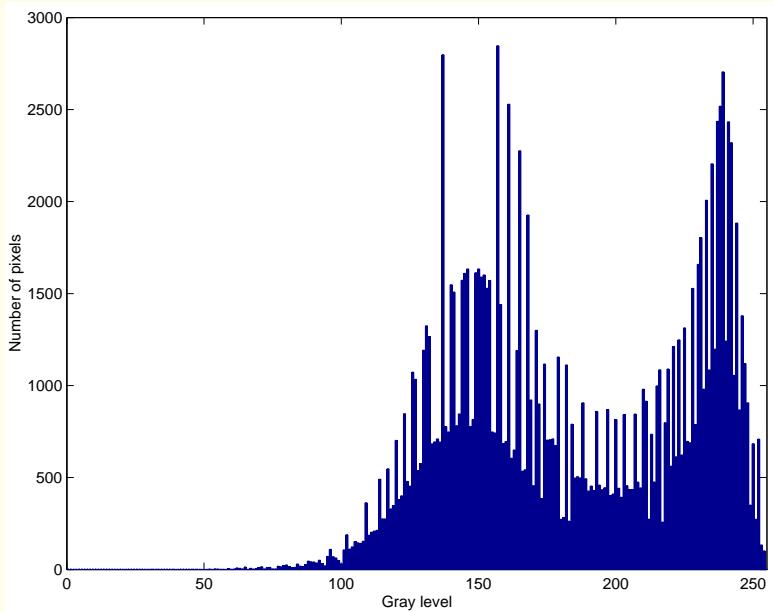
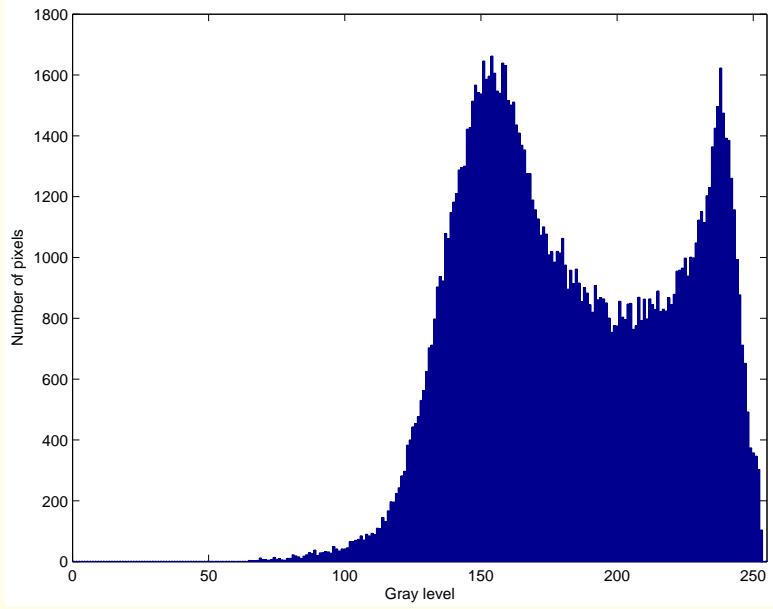


Figure 2.11: Histogram of the image of the ventricular myocyte in Figure 1.3. The size of the image is $480 \times 480 = 230,400$ pixels. Entropy $H = 4.96$ bits.





(a)



(b)

Figure 2.12: (a) Histogram of the image of the collagen fibers in Figure 1.5 (b); $H = 7.0$ bits. (b) Histogram of the image after the application of the 3×3 mean filter and rounding the results to integers; $H = 7.1$ bits.

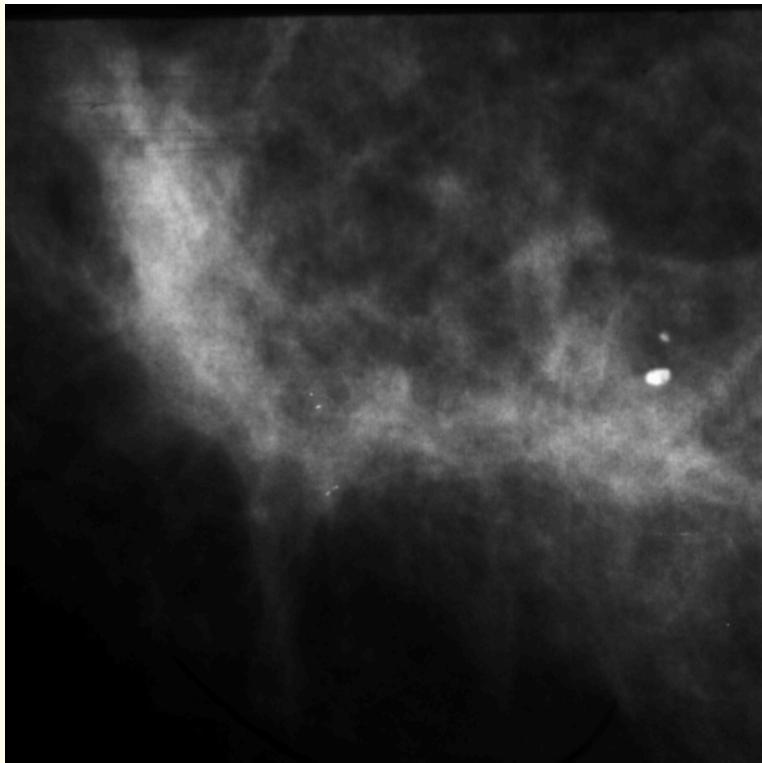


Figure 2.13: Part of a mammogram with a malignant tumor (the relatively bright region along the upper-left edge of the image). The size of the image is $700 \times 700 = 490,000$ pixels. The pixel resolution of $62 \mu\text{m}$; the width of the image is about 44 mm . Image courtesy of Foothills Hospital, Calgary.

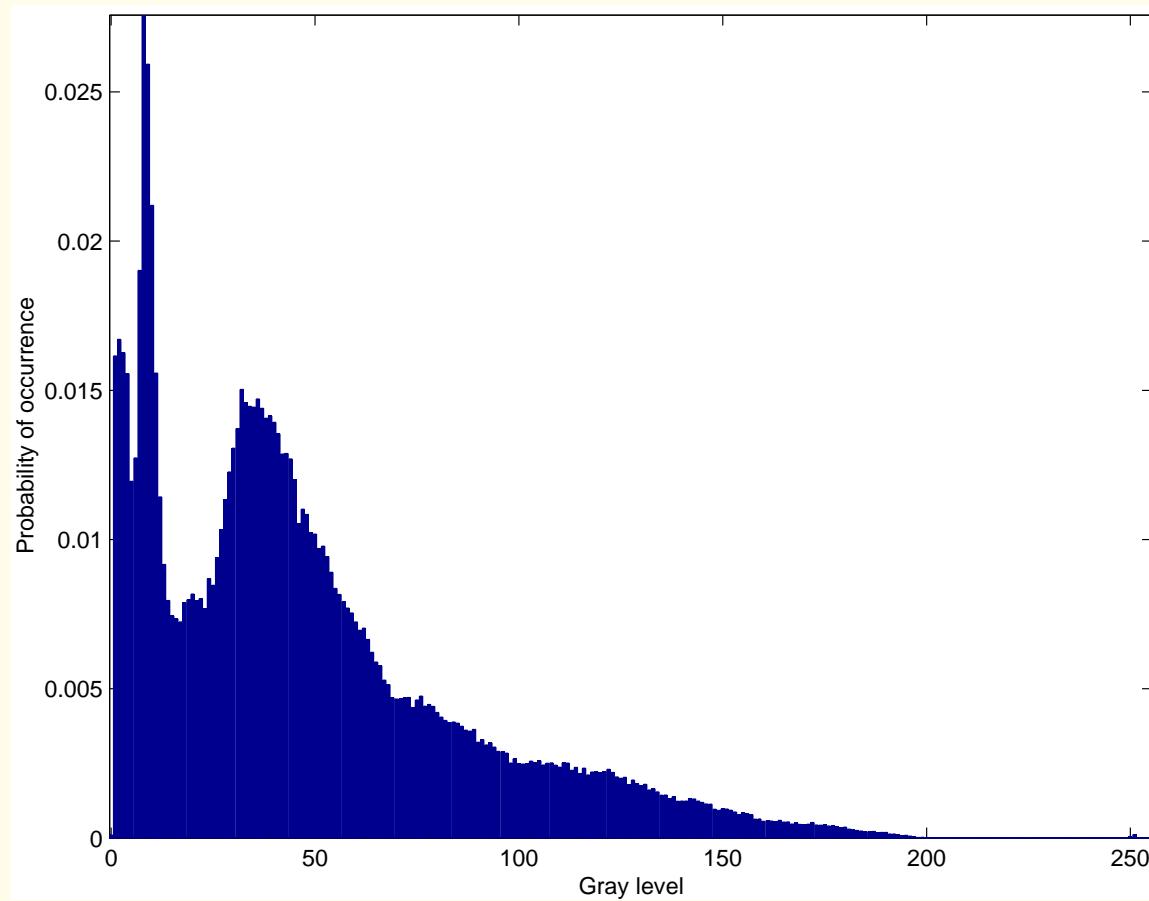


Figure 2.14: Normalized histogram of the mammogram in Figure 2.13. Entropy $H = 6.92 \text{ bits}$.

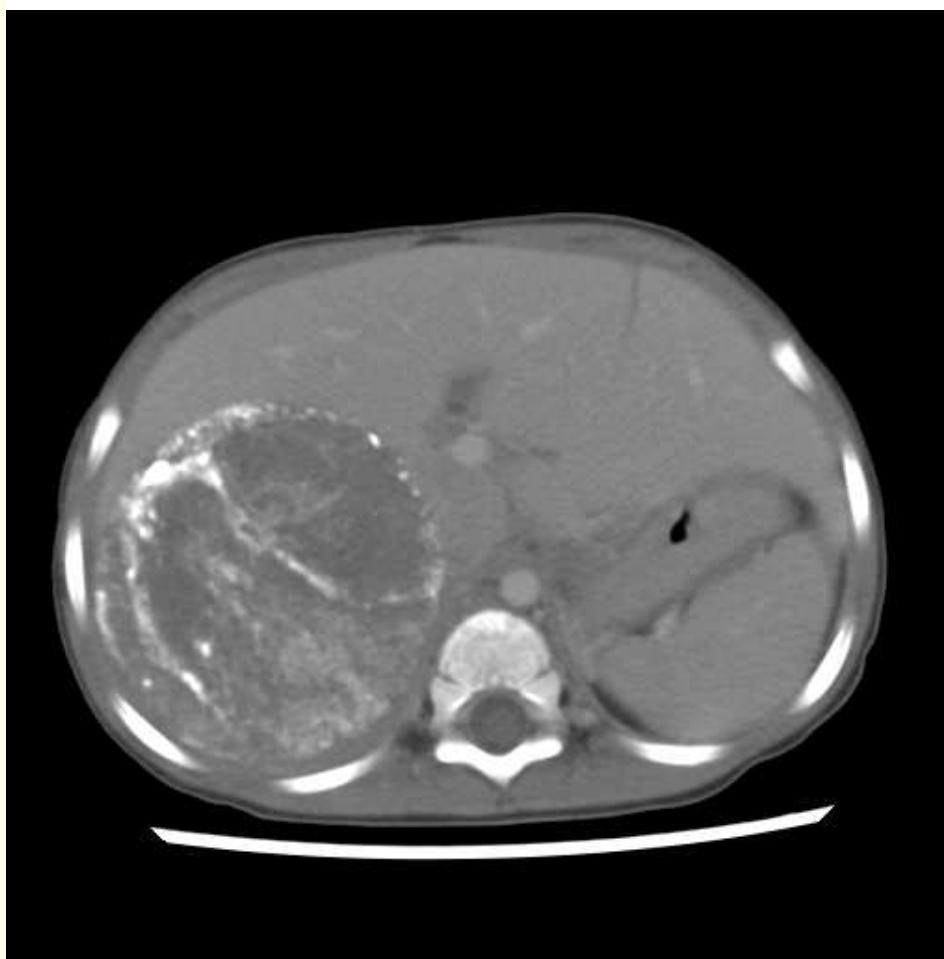
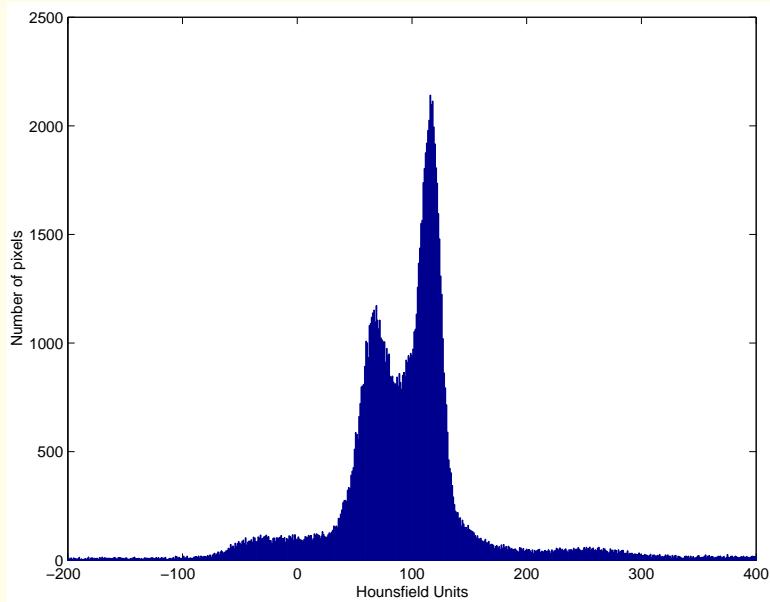
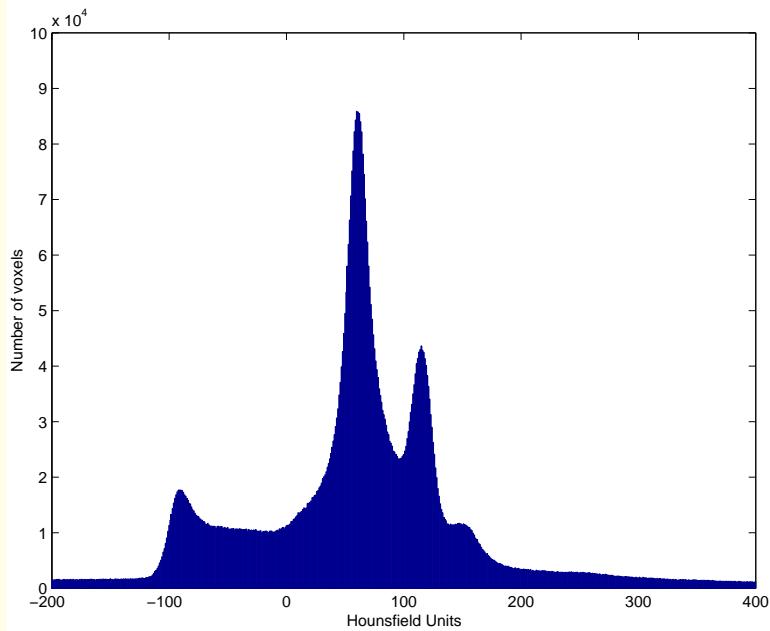


Figure 2.15: CT image of a patient with neuroblastoma. Only one sectional image out of a total of 75 images in the study is shown. The size of the image is $512 \times 512 = 262,144$ pixels. The tumor, which appears as a large circular region on the left-hand side of the image, includes calcified tissues that appear as bright regions. The HU range of $[-200, 400]$ has been linearly mapped to the display range of $[0, 255]$; see also Figures 2.16 and 4.4. Image courtesy of Alberta Children's Hospital, Calgary.



(a)



(b)

Figure 2.16: (a) Histogram of the CT section image in Figure 2.15. (b) Histogram of the entire CT study of the patient, with 75 sectional images. The histograms are displayed for the range $HU = [-200, 400]$ only.



2.8 Entropy

Entropy is a *statistical measure* of information.

Pixels in an image considered to be symbols produced by a discrete information source with the gray levels as its states.

Consider the occurrence of L gray levels in an image,

with the probability of occurrence of the l^{th} gray level being $p(l), l = 0, 1, 2, \dots, L - 1$.

Gray level of a pixel: a random variable.



A measure of information conveyed by an event (a pixel or a gray level) may be related to the *statistical uncertainty* of the event

rather than the semantic or structural content of the image.



A measure of information $h(p)$ should be a function of $p(l)$, satisfying the following criteria:

- $h(p)$ should be continuous for $0 < p < 1$.
- $h(p) = \infty$ for $p = 0$.
- $h(p) = 0$ for $p = 1$.
- $h(p_2) > h(p_1)$ if $p_2 < p_1$.
- If two statistically independent image processes (or pixels) f and g are considered, the joint information of the two sources is the sum of their individual measures of information:
$$h_{f,g} = h_f + h_g.$$

These requirements are met by $h(p) = -\log(p)$.



When a source generates a number of gray levels with different probabilities, a measure of average information or *entropy* is the expected value of information in each possible level:

$$H = \sum_{l=0}^{L-1} p(l) h[p(l)]. \quad (2.17)$$

Using $-\log_2$ in place of h , we obtain

$$H = -\sum_{l=0}^{L-1} p(l) \log_2 [p(l)] \text{ bits.} \quad (2.18)$$

Because the gray levels are considered as individual entities in this definition, that is, no neighboring elements are taken into account, the result is known as the *zeroth-order* entropy.



Differentiating the function in Equation 2.18 with respect to $p(l)$, it can be shown that the maximum possible entropy occurs when all the gray levels occur with the same probability (equal to $\frac{1}{L}$); when the various gray levels are equally likely:

$$H_{\max} = - \sum_{l=0}^{L-1} \frac{1}{L} \log_2 \left[\frac{1}{L} \right] = \log_2 L. \quad (2.19)$$

If the number of gray levels in an image is 2^K , $H_{\max} = K \text{ bits}$.

Maximum possible entropy of an image with 8-bit pixels is 8 bits .



Entropy characterizes the *statistical* information content of a source based upon the PDF of the constituent events, which are treated as random variables.

The measure is not sensitive to the pictorial, structural, semantic, or application-specific (diagnostic) information in the image.

Entropy does not account for the spatial distribution of the gray levels in a given image.

Gives the lower bound on the noise-free transmission rate and storage capacity requirements.



Properties of entropy:

- $H_p \geq 0$, with $H_p = 0$ only for $p = 0$ or $p = 1$:
- The joint information $H_{(p_1, p_2, \dots, p_n)}$ conveyed by n events, with probabilities of occurrence p_1, p_2, \dots, p_n , is governed by
$$H_{(p_1, p_2, \dots, p_n)} \leq \log(n),$$
with equality if and only if $p_i = \frac{1}{n}$ for $i = 1, 2, \dots, n.$



- Considering two images or sources f and g with PDFs $p_f(l_1)$ and $p_g(l_2)$, where l_1 and l_2 represent gray levels in the range $[0, L - 1]$, the average joint information or joint entropy is

$$H_{f,g} = - \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} p_{f,g}(l_1, l_2) \log_2[p_{f,g}(l_1, l_2)]. \quad (2.20)$$

If the two sources are statistically independent, the joint PDF $p_{f,g}(l_1, l_2)$ reduces to $p_f(l_1) p_g(l_2)$.

Joint entropy is governed by the condition

$$H_{f,g} \leq H_f + H_g,$$

with equality iff f and g are statistically independent.



- The conditional entropy of an image f given that another image g has been observed is

$$\begin{aligned} H_{f|g} &= - \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} p_g(l_2) p_{f|g}(l_1, l_2) \log_2 [p_{f|g}(l_1, l_2)] \\ &= - \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} p_{f,g}(l_1, l_2) \log_2 [p_{f|g}(l_1, l_2)], \quad (2.21) \end{aligned}$$

where $p_{f|g}(l_1, l_2)$ is the conditional PDF of f given g .

$$H_{f|g} = H_{f,g} - H_g \leq H_f,$$

with equality iff f and g are statistically independent.



Note: The conditional PDF of f given g is defined as

$$p_{f|g}(l_1, l_2) = \begin{cases} \frac{p_{f,g}(l_1, l_2)}{p_g(l_2)} & \text{if } p_g(l_2) > 0 \\ 1 & \text{otherwise.} \end{cases} \quad (2.22)$$



Higher-order entropy: The definition of the zeroth-order entropy in Equation 2.18 assumes that the successive pixels produced by the source are statistically independent.

While governed by the limit $H_{\max} = K \text{ bits}$, the entropy of a real-world image could be considerably lower:

neighboring pixels are not independent of one another.

It is desirable to consider sequences of pixels to estimate the true entropy or information content of a given image.



Let $p(\{l_n\})$ represent the probability of occurrence of the sequence $\{l_0, l_1, l_2, \dots, l_n\}$ of gray levels in the image f .

n : number of neighboring or additional elements considered, not counting the initial or zeroth element.



The n^{th} -order entropy of f is defined as

$$H_n = -\frac{1}{(n+1)} \sum_{\{l_n\}} p(\{l_n\}) \log_2 [p(\{l_n\})], \quad (2.23)$$

Σ over all possible sequences $\{l_n\}$ with $(n+1)$ pixels.

Variations exist in the definition of higher-order entropy.

H_n is a monotonically decreasing function of n , and approaches the true entropy of the source as $n \rightarrow \infty$.



Mutual information:

Important in analysis of transmission of images over a communication system, as well as storage in and retrieval from an archival system, with potential loss of information.

$$I_{f|g} = H_f + H_g - H_{f,g} = H_f - H_{f|g} = H_g - H_{g|f}. \quad (2.24)$$

Represents the information received or retrieved:

H_f is the information input to the transmission or archival system in the form of the image f .

$H_{f|g}$ is the information about f given that the received or retrieved image g has been observed.



If g is completely correlated with f

$$H_{f|g} = 0 \text{ and } I_{f|g} = H_f$$

no loss or distortion in image transmission and reception.

If g is independent of f

$$H_{f|g} = H_f \text{ and } I_{f|g} = 0$$

complete loss of information in transmission or archival.

Entropy and mutual information useful in the design and analysis of image archival, coding, and communication systems.



2.9 Blur and Spread Functions

Several components of image acquisition systems cause blurring due to intrinsic and practical limitations.

The simplest visualization of blurring is provided by using a single, ideal point to represent the object being imaged.

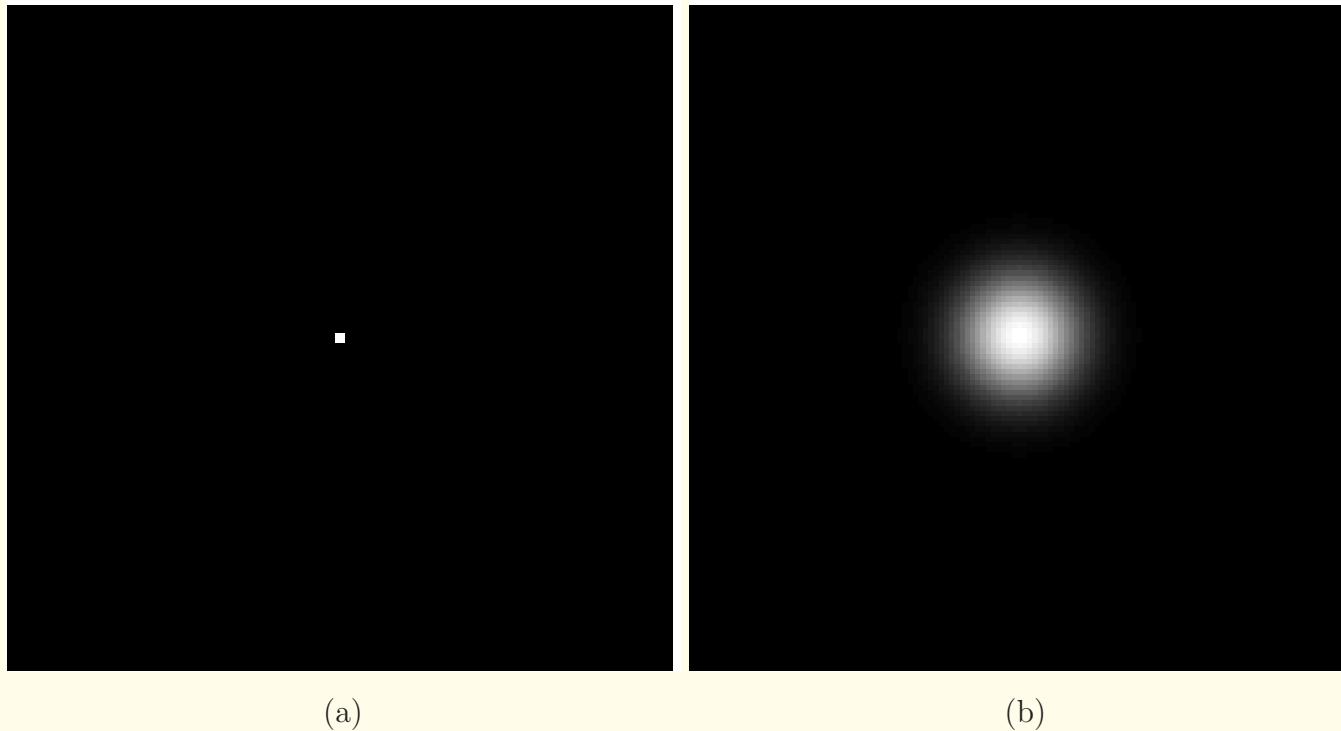


Figure 2.17: (a) An ideal point source. (b) A Gaussian-shaped point spread function.



Mathematically, an ideal point is represented by the continuous unit impulse function or the Dirac delta function $\delta(x, y)$:

$$\delta(x, y) = \begin{cases} \text{undefined at } x = 0, y = 0 \\ 0 \quad \text{otherwise,} \end{cases} \quad (2.25)$$

and

$$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} \delta(x, y) dx dy = 1. \quad (2.26)$$



The 1D Dirac delta function:

$\delta(x)$ is defined in terms of its action within an integral as

$$\int_a^b f(x) \delta(x - x_o) dx = \begin{cases} f(x_o) & \text{if } a < x_o < b \\ 0 & \text{otherwise,} \end{cases} \quad (2.27)$$

where $f(x)$ is a function that is continuous at x_o .

This is known as the *sifting property*.



The expression may be extended to all x as

$$f(x) = \int_{\alpha=-\infty}^{\infty} f(\alpha) \delta(x - \alpha) d\alpha. \quad (2.28)$$

Resolving the arbitrary signal $f(x)$ into a weighted combination of mutually orthogonal delta functions.



Delta function in terms of its integrated strength:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1, \quad (2.29)$$

with the conditions

$$\delta(x) = \begin{cases} \text{undefined at } x = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (2.30)$$



The delta function is also defined as the limiting condition of several ordinary functions, one of which is

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \exp\left(-\frac{|x|}{\epsilon}\right). \quad (2.31)$$

The delta function may be visualized as the limit of a function with a sharp peak of undefined value, whose integral over the full extent of the independent variable is maintained as unity while its temporal or spatial extent is compressed toward zero.



The image obtained when the input is a point or impulse function is known as the impulse response or

point spread function (PSF).



Assuming the imaging system to be linear and shift-invariant (LSI) (or position-invariant or space-invariant), the image $g(x, y)$ of an object $f(x, y)$ is given by the 2D convolution integral

$$\begin{aligned} g(x, y) &= \int_{\alpha=-\infty}^{\infty} \int_{\beta=-\infty}^{\infty} h(x - \alpha, y - \beta) f(\alpha, \beta) d\alpha d\beta \\ &= \int_{\alpha=-\infty}^{\infty} \int_{\beta=-\infty}^{\infty} h(\alpha, \beta) f(x - \alpha, y - \beta) d\alpha d\beta \\ &= h(x, y) * f(x, y), \end{aligned} \tag{2.32}$$

where $h(x, y)$ is the PSF, α and β are temporary variables of integration, and $*$ represents 2D convolution.



Some examples of the cause of blurring are:

- Focal spot.
- Thickness of screen or crystal.
- Scattering.

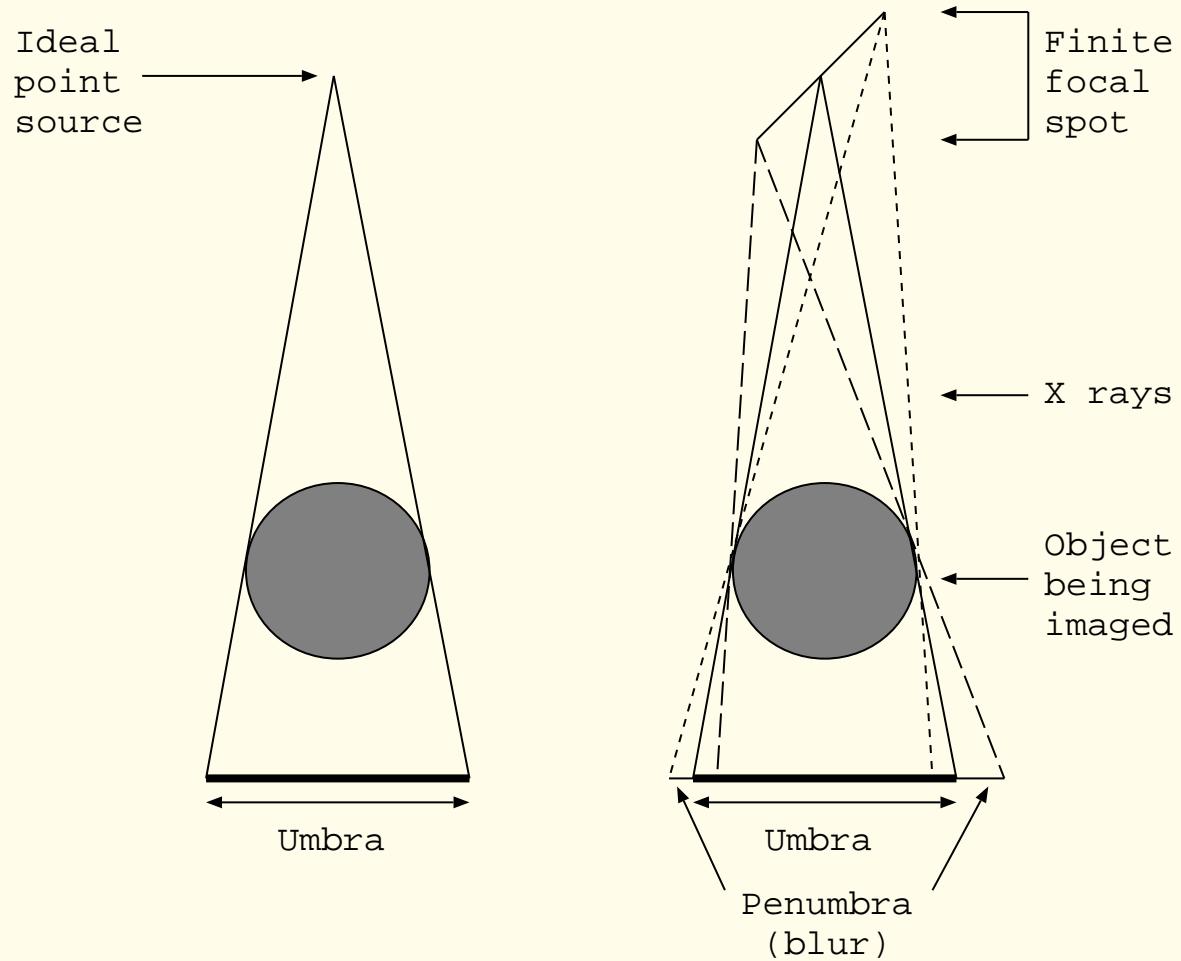


Figure 2.18: The effect of a finite focal spot (X-ray-generating portion of the target) on the sharpness of the image of an object.



Point, line, and edge spread functions:

It is often not possible to obtain an image of an ideal point.

However, it is possible to construct phantoms to represent ideal lines or edges.

An image obtained of line function is known as the *line spread function* (LSF) of the system.

A cross-section of an ideal straight line is a point (impulse).

The reconstruction of a cross-section of a line phantom provides the PSF of the system.

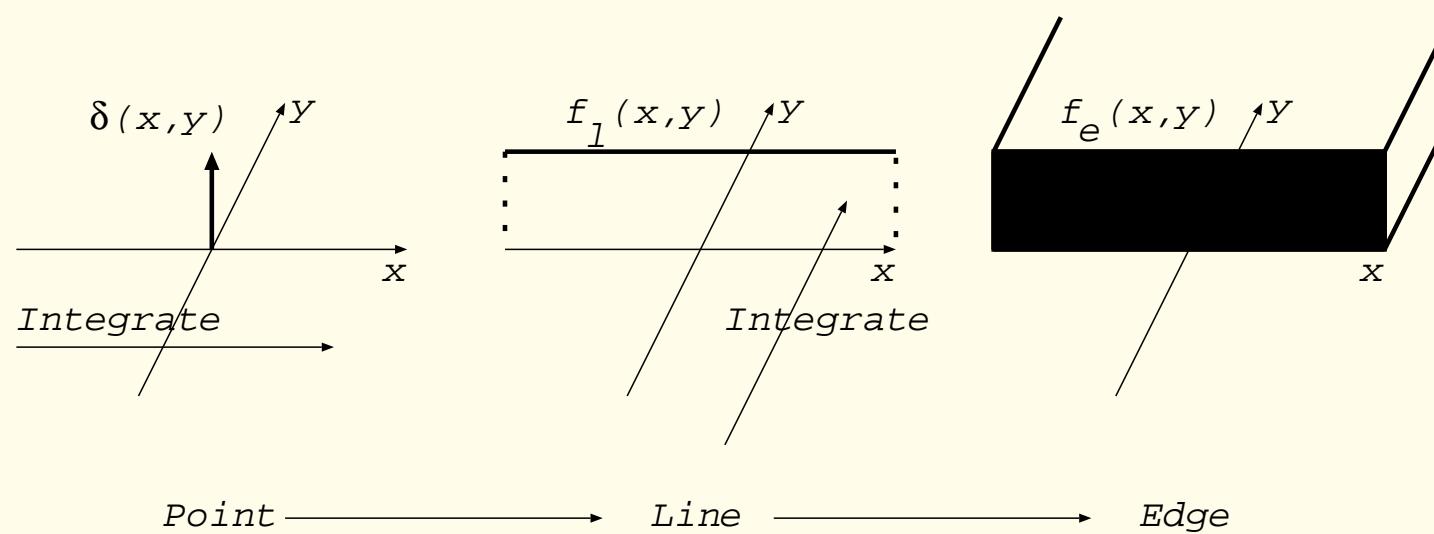


Figure 2.19: The relationship between point (impulse function), line, and edge (step) images. The height of each function represents its strength.



A phantom representing an ideal edge may also be used.

A profile of the image of such a phantom across the ideal edge provides the *edge spread function* (ESF).

The derivative of the ESF gives the LSF of the system.

The PSF may be estimated from the LSF.

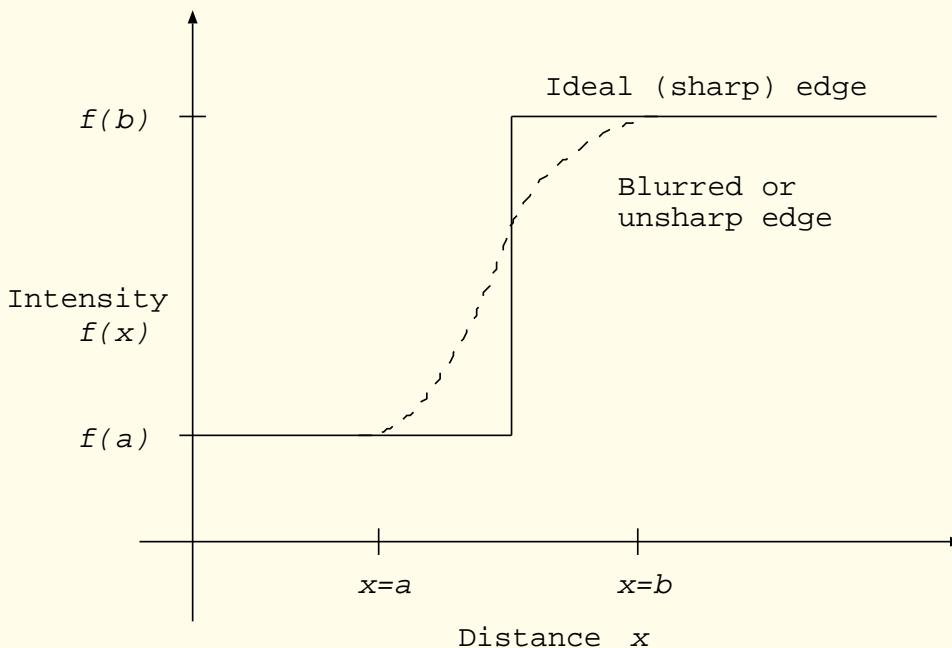


Figure 2.20: Blurring of an ideal sharp edge into an unsharp edge by an imaging system.



Mathematical relationships between the PSF, LSF, and ESF:

Consider integration of the 2D delta function along the x axis:

$$\begin{aligned} f_l(x, y) &= \int_{x=-\infty}^{\infty} \delta(x, y) dx \\ &= \int_{x=-\infty}^{\infty} \delta(x) \delta(y) dx \\ &= \delta(y) \int_{x=-\infty}^{\infty} \delta(x) dx \\ &= \delta(y). \end{aligned} \tag{2.33}$$



The last integral above is equal to unity.

The separability property of the 2D impulse function as $\delta(x, y) = \delta(x) \delta(y)$ has been used above.

$\delta(y)$ over the 2D (x, y) space is a line function on the x axis.



The output of an LSI system when the input is the line image $f_l(x, y) = \delta(y)$, that is, the LSF $h_l(x, y)$, is

$$\begin{aligned} h_l(x, y) &= \int_{\alpha=-\infty}^{\infty} \int_{\beta=-\infty}^{\infty} h(\alpha, \beta) f_l(x - \alpha, y - \beta) d\alpha d\beta \\ &= \int_{\alpha=-\infty}^{\infty} \int_{\beta=-\infty}^{\infty} h(\alpha, \beta) \delta(y - \beta) d\alpha d\beta \\ &= \int_{\alpha=-\infty}^{\infty} h(\alpha, y) d\alpha \\ &= \int_{x=-\infty}^{\infty} h(x, y) dx. \end{aligned} \tag{2.34}$$

$h(x, y)$ is the PSF; the LSF is the integral of the PSF.



Consider the Fourier transform of $h_l(x, y)$.

$$\begin{aligned} H_l(v) &= \int_{y=-\infty}^{\infty} h_l(y) \exp(-j2\pi vy) dy \\ &= \int_{y=-\infty}^{\infty} dy \int_{x=-\infty}^{\infty} dx h(x, y) \exp[-j2\pi(ux + vy)]|_{u=0} \\ &= H(u, v)|_{u=0} \\ &= H(0, v), \end{aligned} \tag{2.35}$$



$H(u, v)$ is the 2D Fourier transform of $h(x, y)$:

the *modulation transfer function (MTF)*.

The Fourier transform of the LSF gives the values of the Fourier transform of the PSF along a line in the 2D Fourier plane (in this case, along the v axis).



Consider integrating the line function as follows:

$$\begin{aligned} f_e(x, y) &= \int_{\beta=-\infty}^y f_l(x, \beta) d\beta \\ &= \int_{\beta=-\infty}^y \delta(\beta) d\beta. \end{aligned} \quad (2.36)$$



The resulting function has the property

$$\forall x, f_e(x, y) = \begin{cases} 1 & \text{if } y > 0 \\ 0 & \text{if } y < 0, \end{cases} \quad (2.37)$$

which is an edge or unit step function parallel to the x axis.

Thus, the edge or step function is obtained by integrating the line function.



The ESF is given by

$$h_e(y) = \int_{\beta=-\infty}^y h_l(\beta) d\beta. \quad (2.38)$$

Conversely, the LSF is the derivative of the ESF:

$$h_l(y) = \frac{d}{dy} h_e(y). \quad (2.39)$$

Thus the ESF may be used to obtain the LSF, which may further be used to obtain the PSF and MTF.



From the Fourier slice theorem:

The Fourier transform of a profile of the LSF is equal to the radial profile of the Fourier transform of the PSF at the angle of placement of the line source.

If the imaging system may be assumed to be isotropic in the plane of the line source, a single radial profile is adequate to reconstruct the complete 2D Fourier transform of the PSF.

An inverse 2D Fourier transform provides the PSF.



Line source prepared using a plastic tube of internal radius 1 mm, filled with 1 mCi (milli Curie) of ^{99m}Tc .

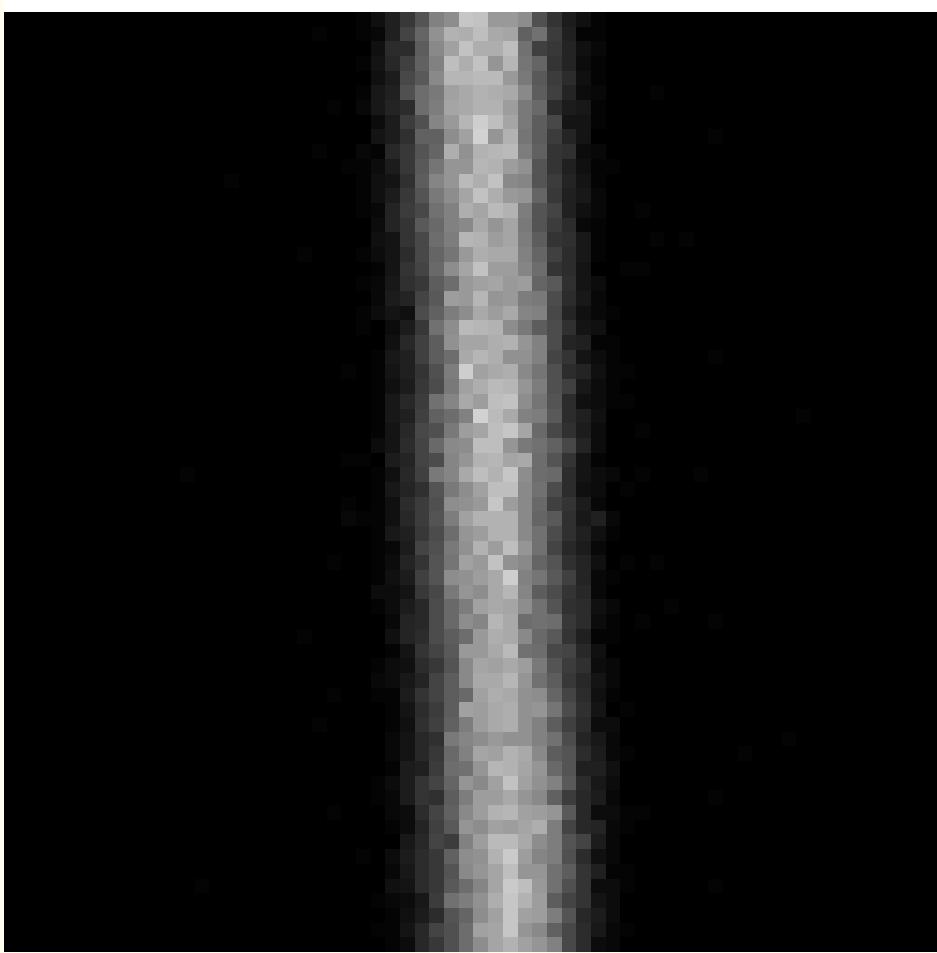


Figure 2.21: Nuclear medicine (planar) image of a line source obtained using a gamma camera. The size of the image is 64×64 pixels, with an effective width of 100 mm. The pixel size is 1.56 mm.

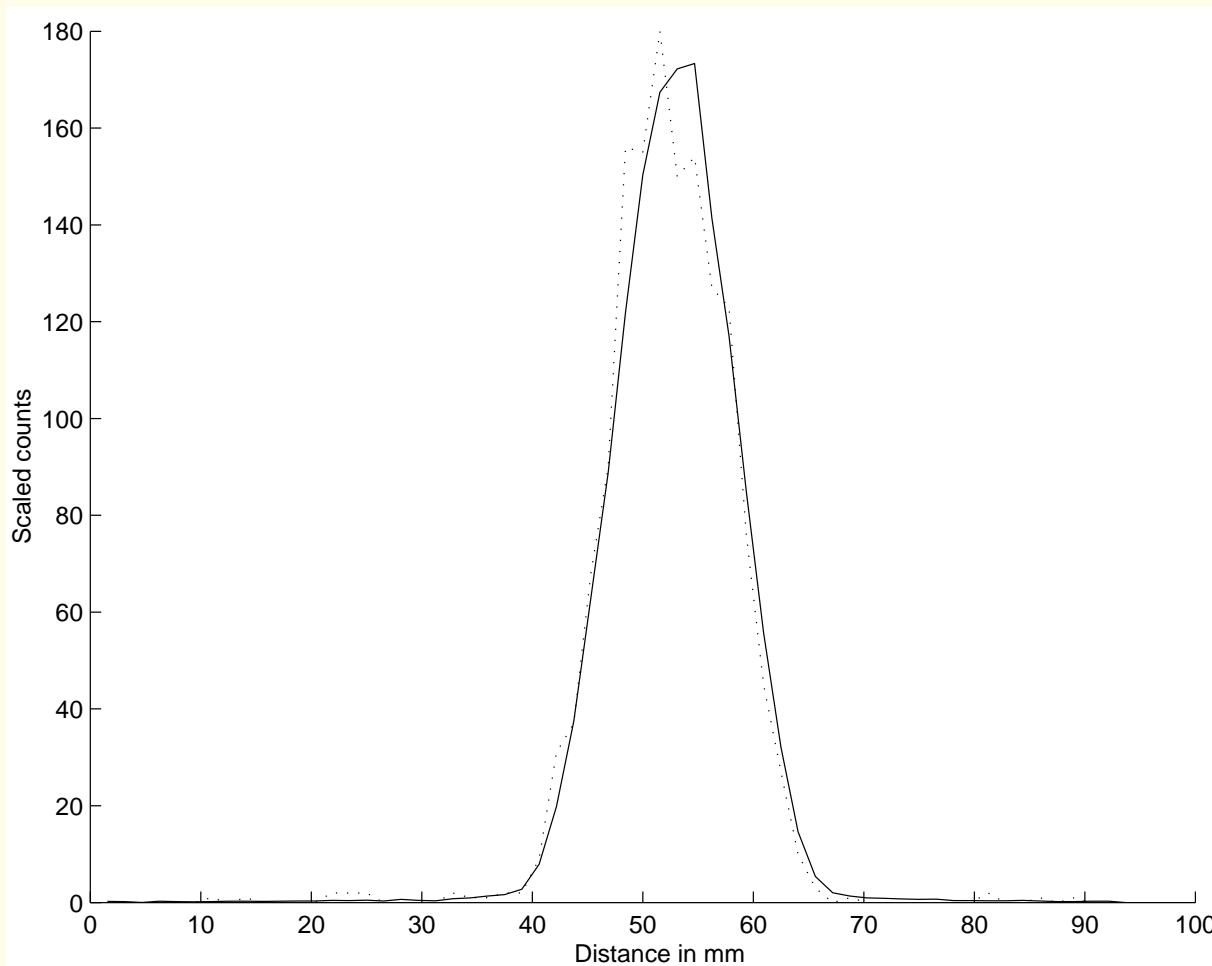


Figure 2.22: Sample profile (dotted line) and averaged profile (solid line) obtained from the image in Figure 2.21. Either profile may be taken to represent the LSF of the gamma camera.

Full width at half the maximum (FWHM): $0.5 - 1.7 \text{ cm}$.



2.10 Resolution

The spatial resolution of an imaging system or an image may be expressed in terms of:

- The sampling interval (in, for example, mm or μm).
- The width of (a profile of) the PSF, usually FWHM (in mm).
- The size of the laser spot used to obtain the digital image by scanning an original film, or the size of the solid-state detector used to obtain the digital image (in μm).
- The smallest visible object or separation between objects in the image (in mm or μm).
- The finest grid pattern that remains visible in the image (in lp/mm).



Typical resolution limits of a few imaging systems:

- X-ray film: $25 - 100 \text{ lp/mm}$.
- screen-film combination: $5 - 10 \text{ lp/mm}$;
mammography: up to 20 lp/mm .
- CT: 0.7 lp/mm ;
 μCT : 50 lp/mm or $10 \mu\text{m}$;
- SPECT: $< 0.1 \text{ lp/mm}$.



2.11 The Fourier Transform and Spectral Content

The Fourier transform is a linear, reversible transform that maps an image from the space domain to the frequency domain.

Converting an image from the spatial to the frequency (Fourier) domain helps in

- assessing the spectral content,
- assessing the energy distribution over frequency bands,
- designing filters to remove noise,
- designing filters to enhance the image,
- extracting certain components that are better separated in the frequency domain than in the space domain.



2D Fourier transform of an image $f(x, y)$ is denoted by $F(u, v)$:

$$F(u, v) = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x, y) \exp[-j 2\pi(ux+vy)] dx dy. \quad (2.40)$$

u, v : frequency in the horizontal and vertical directions.



$$\exp[-j 2\pi (ux + vy)] \quad (2.41)$$

$$= \exp(-j 2\pi ux) \exp(-j 2\pi vy)$$

$$= [\cos(2\pi ux) - j \sin(2\pi ux)] [\cos(2\pi vy) - j \sin(2\pi vy)].$$



Images are typically functions of space.

Units of measurement in the image domain:

m , cm , mm , μm , etc.

In the 2D Fourier domain, the unit of frequency is
 $cycles/mm$, $cycles/m$, mm^{-1} , etc.

Frequency is also expressed as lp/mm .

If the distance to the viewer is taken into account, frequency could be expressed in terms of $cycles/degree$ of the visual angle subtended at the viewer's eye.

The unit *Hertz* is not used in 2D Fourier analysis.



It is common to use the discrete Fourier transform (DFT) via the fast Fourier transform (FFT) algorithm.

2D DFT of a digital image $f(m, n)$ of size $M \times N$ pixels:

$$F(k, l) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \exp \left[-j 2\pi \left(\frac{mk}{M} + \frac{nl}{N} \right) \right]. \quad (2.42)$$

For complete recovery of $f(m, n)$ from $F(k, l)$, the latter should be computed for $k = 0, 1, \dots, M - 1$, and $l = 0, 1, \dots, N - 1$, at the minimum.



Then, the inverse transform gives back the original image with no error or loss of information as

$$f(m, n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l) \exp \left[+j 2\pi \left(\frac{mk}{M} + \frac{nl}{N} \right) \right], \quad (2.43)$$

for $m = 0, 1, \dots, M - 1$, and $n = 0, 1, \dots, N - 1$.

This expression may be interpreted as resolving the given image into a weighted sum of mutually orthogonal exponential (sinusoidal) basis functions.

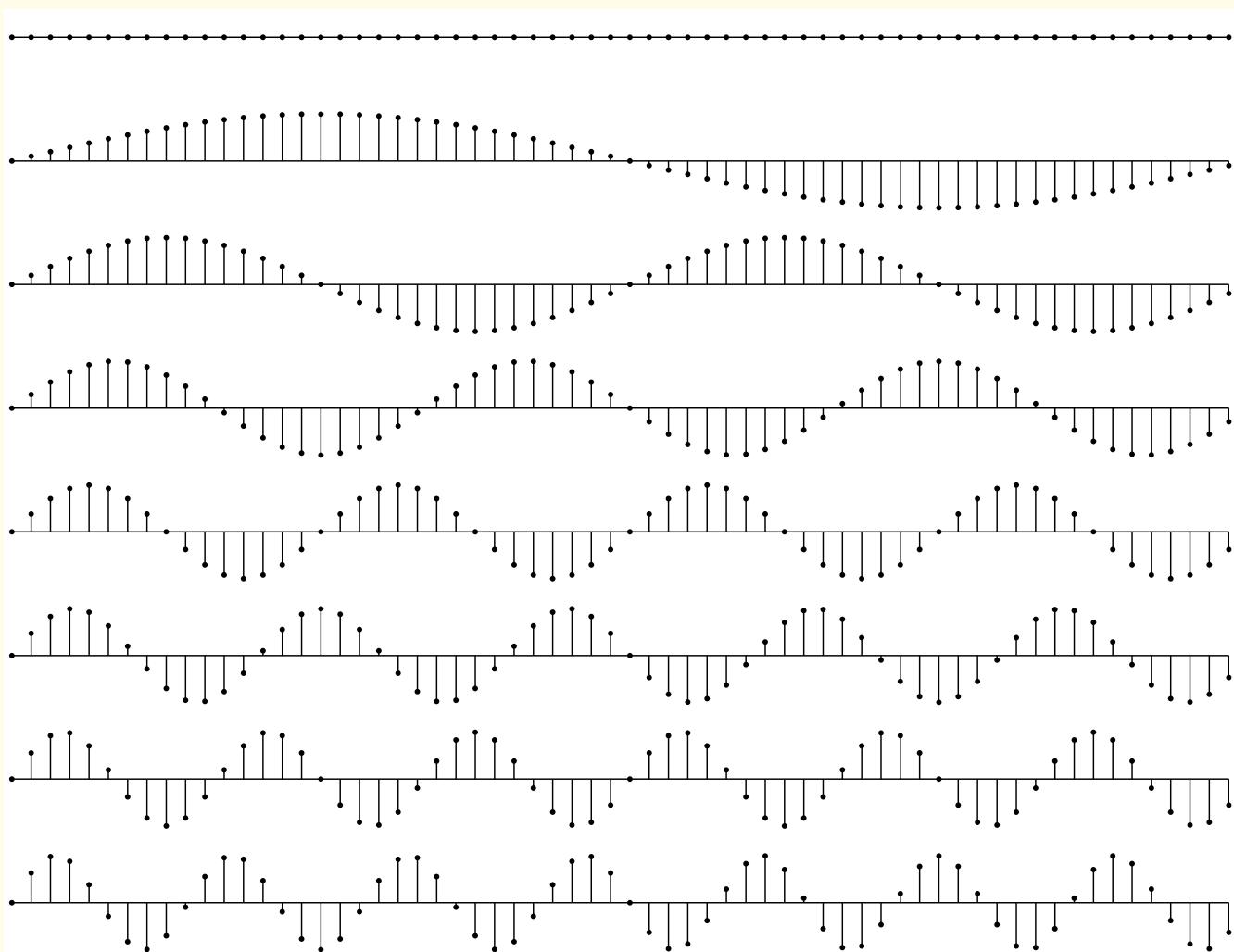


Figure 2.23: The first eight sine basis functions of the 1D DFT; $k = 0, 1, 2, \dots, 7$ from top to bottom. Each function was computed using 64 samples.

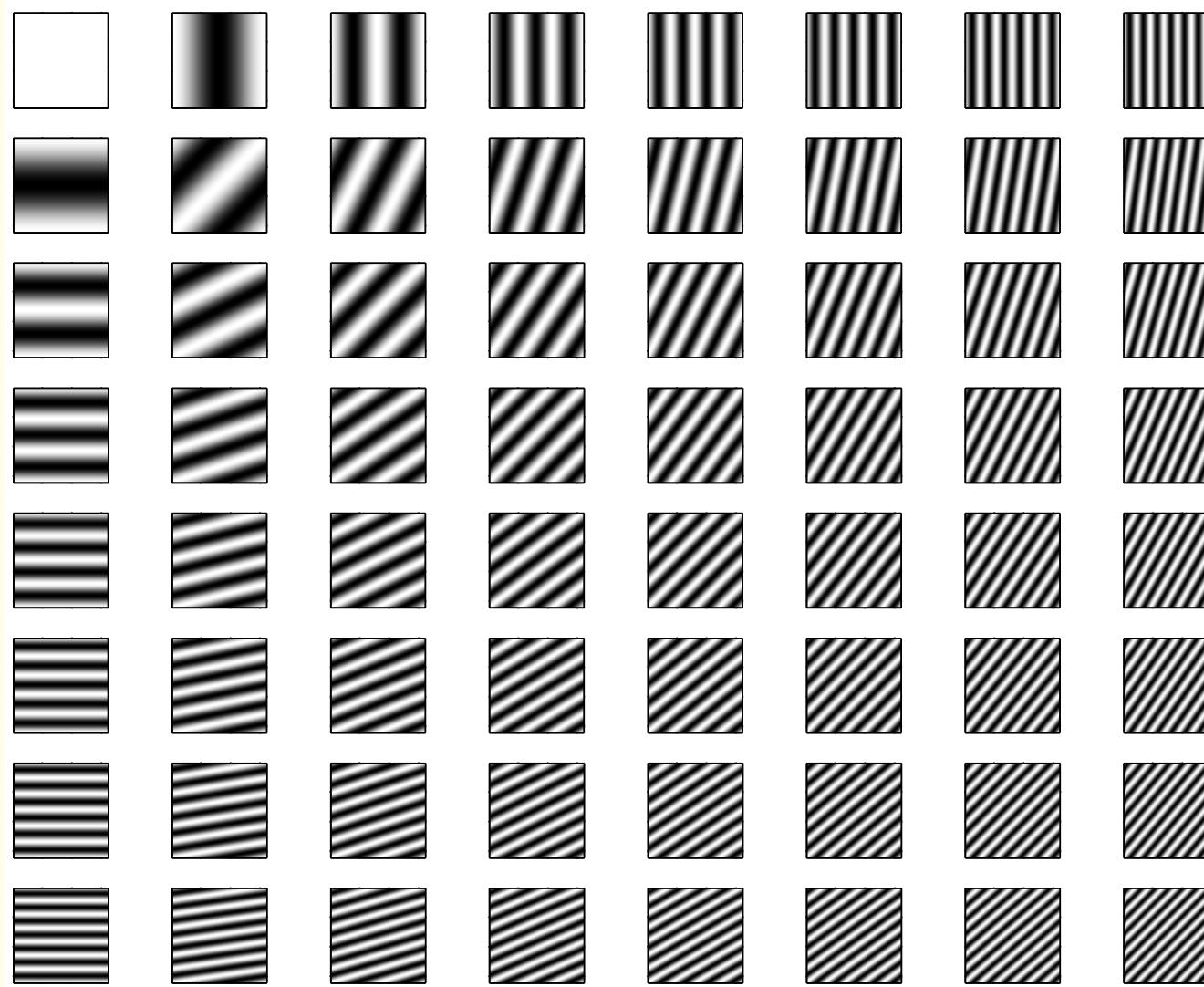


Figure 2.24: The first 64 cosine basis functions of the 2D DFT. Each function was computed using a 64×64 matrix.

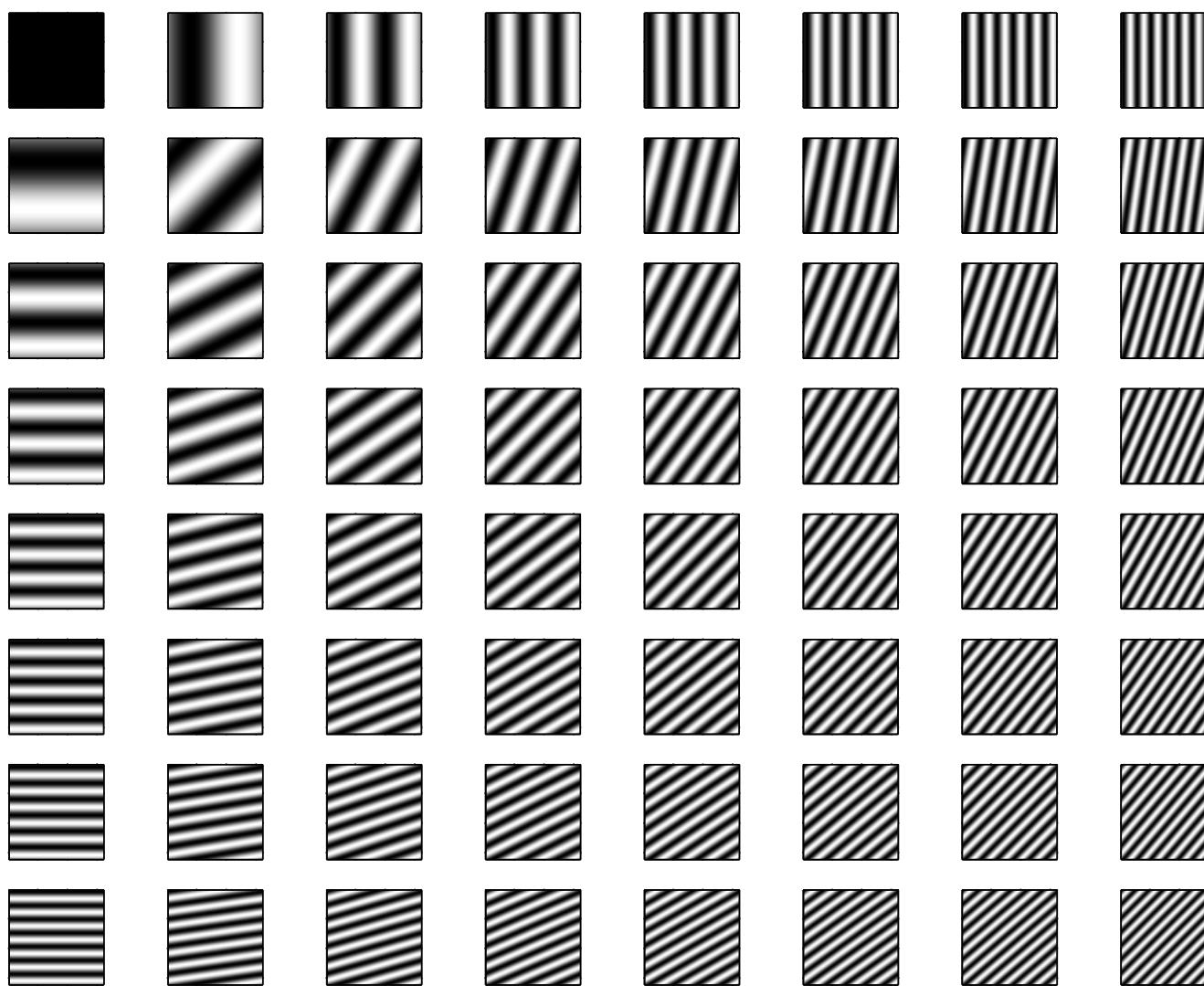


Figure 2.25: The first 64 sine basis functions of the 2D DFT. Each function was computed using a 64×64 matrix.



Zero padding for FFT: Pad the given image with zeros or some other appropriate background value and convert the image to a square of size $N \times N$ where N is an integral power of 2.

Then, all indices in the DFT run from 0 to $N - 1$:



$$F(k, l) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \exp \left[-j \frac{2\pi}{N} (mk + nl) \right], \quad (2.44)$$

with $k = 0, 1, \dots, N - 1$, and $l = 0, 1, \dots, N - 1$.

$$f(m, n) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k, l) \exp \left[+j \frac{2\pi}{N} (mk + nl) \right]. \quad (2.45)$$

The normalization factor has been divided equally between the forward and inverse transforms to be $\frac{1}{N}$ for the sake of symmetry.



The rectangle function and its Fourier transform:

2D function with a rectangular base of size $X \times Y$ and height A :

$$\begin{aligned} f(x, y) &= A \quad \text{if } 0 \leq x \leq X; 0 \leq y \leq Y \quad (2.46) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$



$$\begin{aligned} F(u, v) = & AXY \left[\frac{\sin(\pi u X)}{\pi u X} \exp(-j\pi u X) \right] \\ & \times \left[\frac{\sin(\pi v Y)}{\pi v Y} \exp(-j\pi v Y) \right]. \end{aligned} \quad (2.47)$$

The Fourier transform of a real image is, in general,
a complex function.

The $\exp[]$ functions in Equation 2.47 indicate the
phase components of the spectrum.

An image with even symmetry about the origin will have a
real Fourier transform.



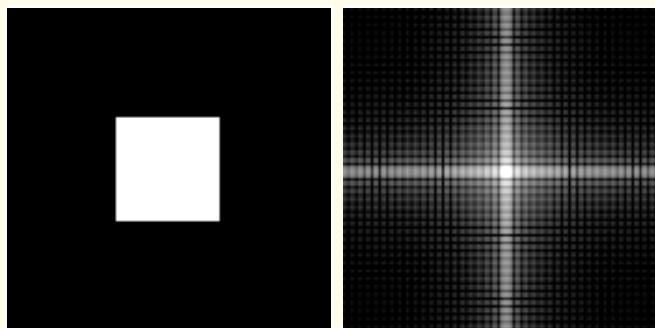
$$\text{rect}(x, y) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2}, |y| < \frac{1}{2} \\ 0 & \text{if } |x| > \frac{1}{2}, |y| > \frac{1}{2}. \end{cases} \quad (2.48)$$

The Fourier transform of the rect function is the sinc function:

$$\text{rect}(x, y) \Leftrightarrow \text{sinc}(u, v). \quad (2.49)$$

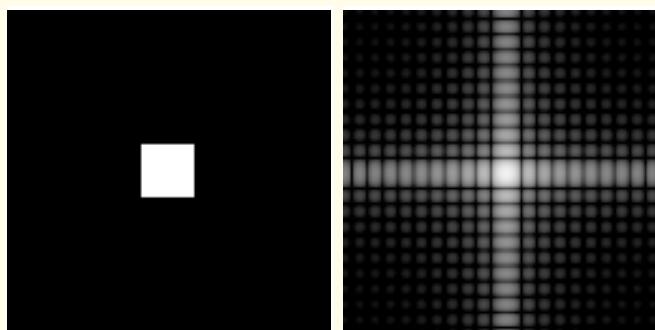
\Leftrightarrow forward and inverse Fourier-transform pair.

$$\text{sinc}(u, v) = \text{sinc}(u) \text{sinc}(v) = \frac{\sin(\pi u)}{\pi u} \frac{\sin(\pi v)}{\pi v}. \quad (2.50)$$



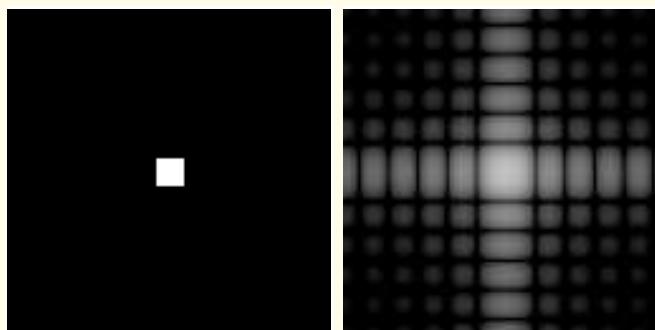
(a)

(b)



(c)

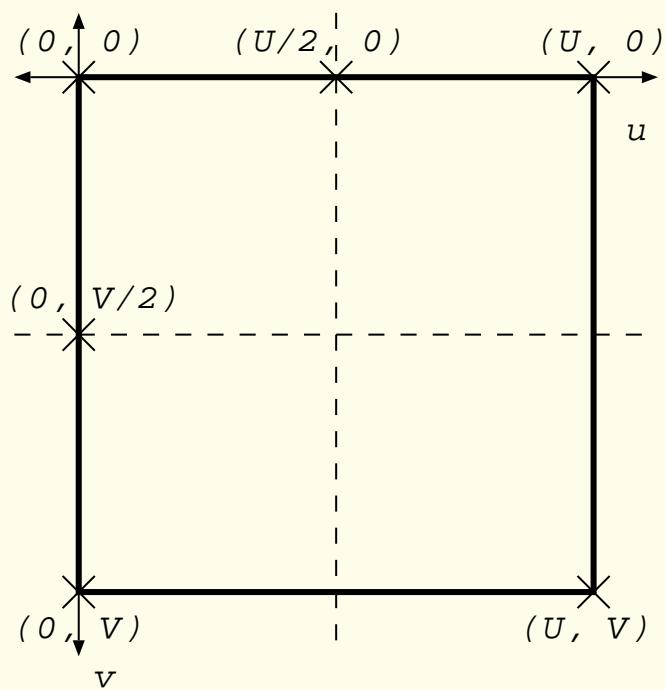
(d)



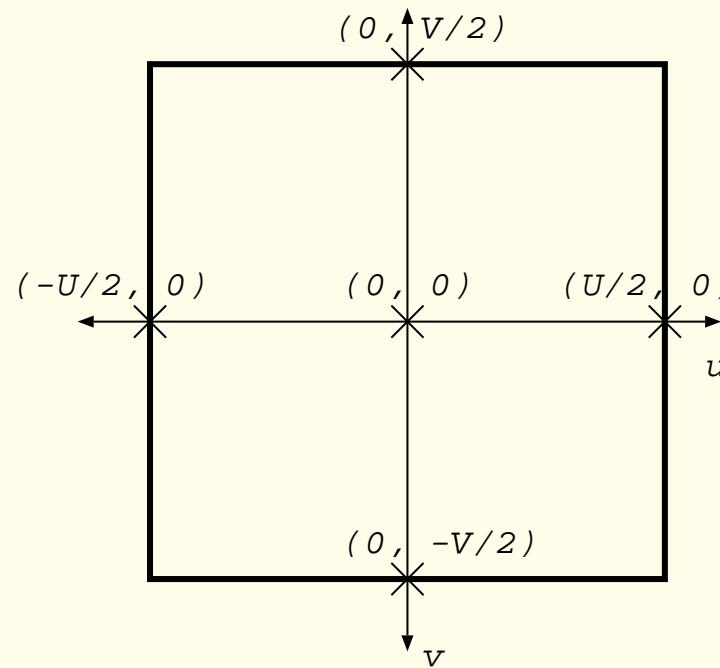
(e)

(f)

Figure 2.26: (a) Rectangle image, with total size 128×128 pixels and a rectangle (square) of size 40×40 pixels. (b) Log-magnitude spectrum of the image in (a). (c) Rectangle size 20×20 pixels. (d) Log-magnitude spectrum of the image in (c). (e) Rectangle size 10×10 pixels. (f) Log-magnitude spectrum of the image in (e). The spectra have been scaled to map the range $[5, 12]$ to the display range $[0, 255]$. See also Figures 2.28 and 2.29.



(a)



(b)

Figure 2.27: Frequency coordinates in (a) the unshifted mode and (b) the shifted mode of display of image spectra. U and V represent the sampling frequencies along the two axes. Spectra of images with real values possess conjugate symmetry about $U/2$ and $V/2$. Spectra of sampled images are periodic, with the periods equal to U and V along the two axes. It is common practice to display one complete period of the shifted spectrum, including the conjugate symmetric parts, as in (b). See also Figure 2.28.

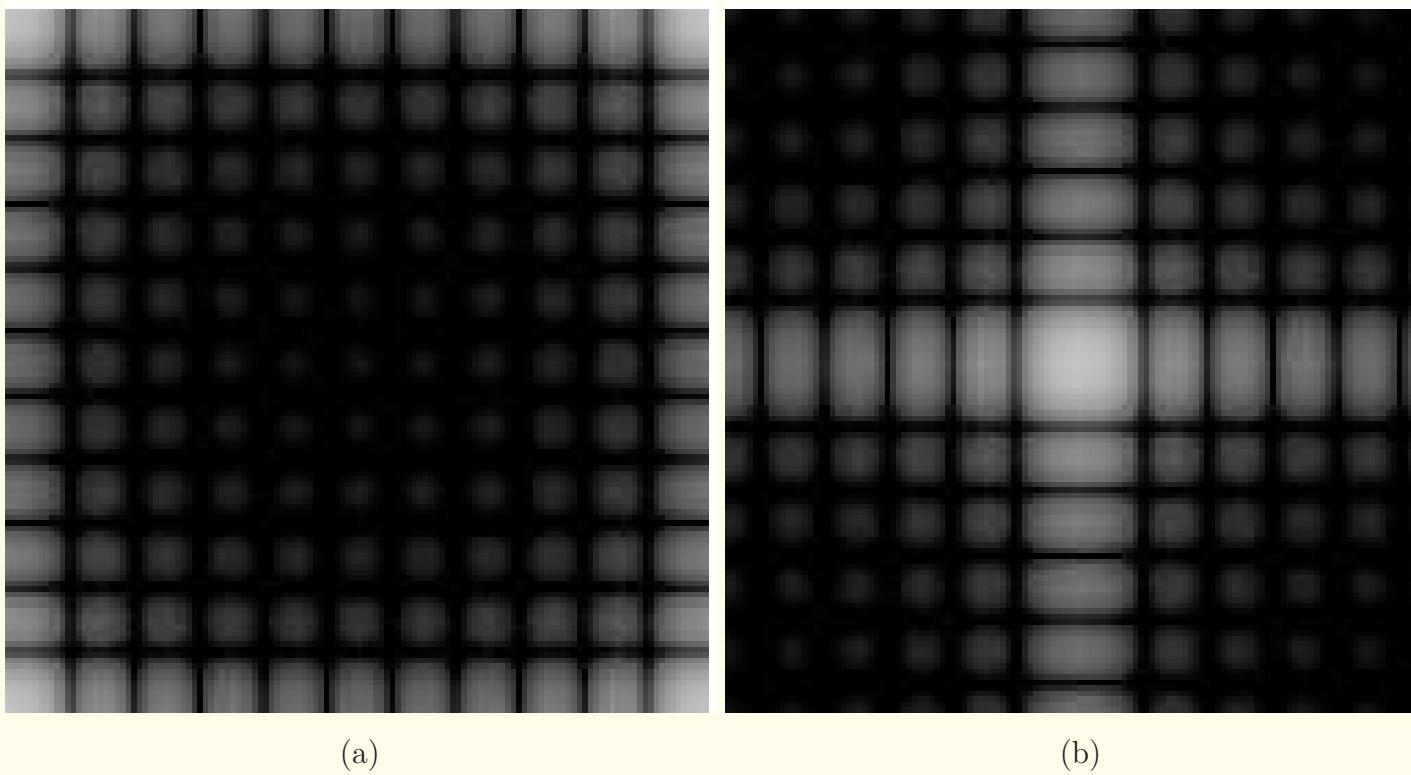
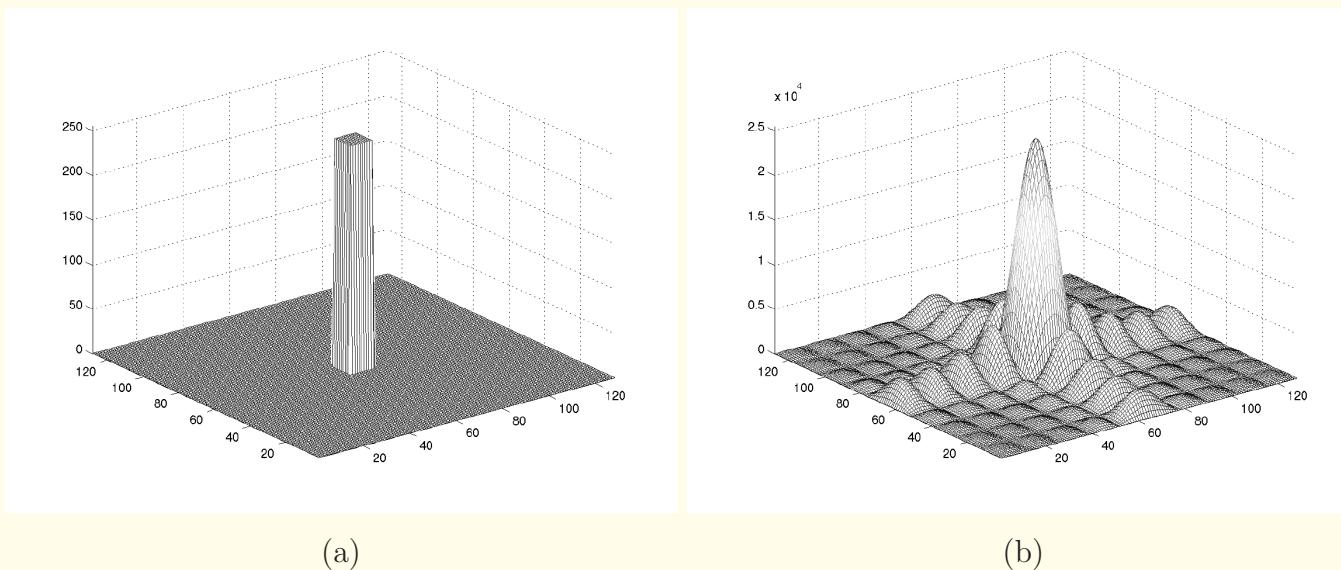


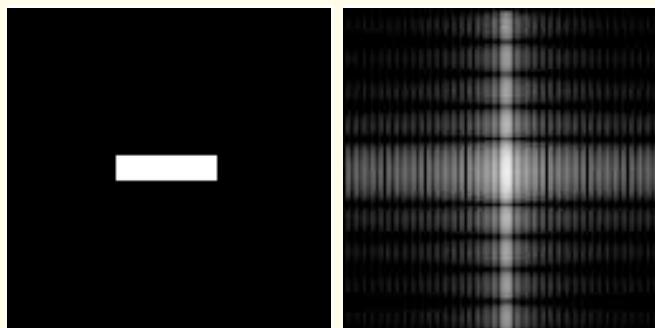
Figure 2.28: (a) Log-magnitude spectrum of the rectangle image in Figure 2.26 (e) without shifting. Most FFT routines provide spectral data in this format. (b) The spectrum in (a) shifted or folded such that $(u, v) = (0, 0)$ is at the center. It is common practice to display one complete period of the shifted spectrum, including the conjugate symmetric parts, as in (b). See also Figure 2.27.



(a)

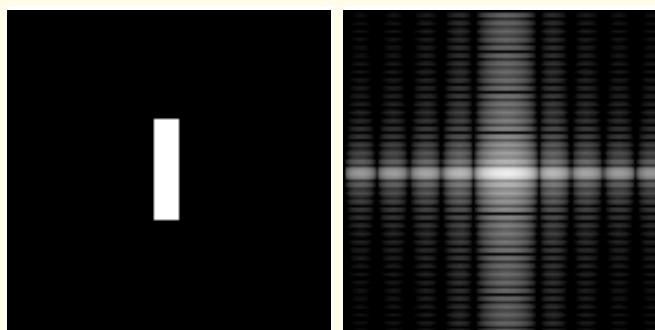
(b)

Figure 2.29: (a) Mesh plot of the rectangle image in Figure 2.26 (e), with total size 128×128 pixels and a rectangle (square) of size 10×10 pixels. (b) Magnitude spectrum of the image in (a).



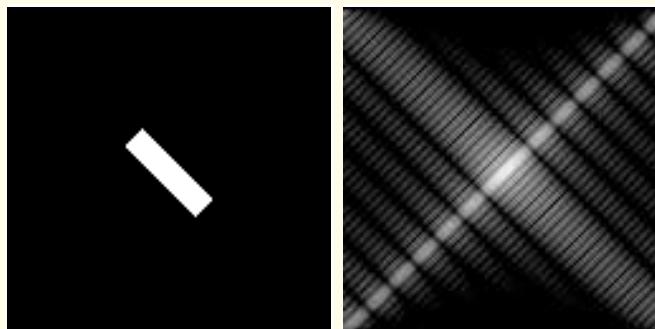
(a)

(b)



(c)

(d)



(e)

(f)

Figure 2.30: (a) Rectangle image, with total size 128×128 pixels and a rectangle of size 10×40 pixels. (b) Log-magnitude spectrum of the image in (a). (c) Rectangle size 40×10 pixels; this image may be considered to be that in (a) rotated by 90° . (d) Log-magnitude spectrum of the image in (c). (e) Image in (c) rotated by 45° using nearest-neighbor selection. (f) Log-magnitude spectrum of the image in (e). Spectra scaled to map $[5, 12]$ to the display range $[0, 255]$.



The circle function and its Fourier transform:

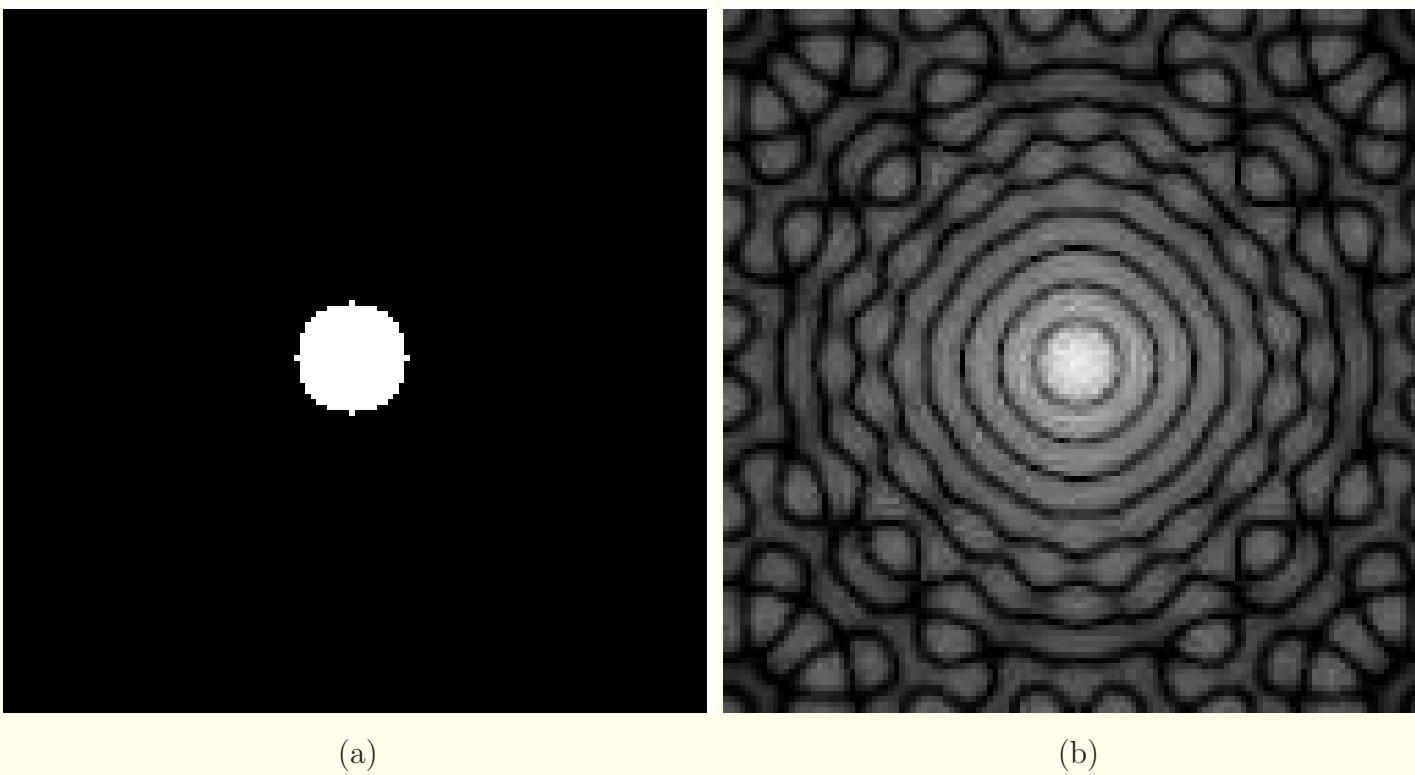
Circular apertures and functions are encountered often in imaging and image processing.

$$\text{circ}(r) = \begin{cases} 1 & \text{if } r < 1 \\ 0 & \text{if } r > 1, \end{cases} \quad (2.51)$$

where $r = \sqrt{(x^2 + y^2)}$.

The Fourier transform of $\text{circ}(r)$ is $\frac{1}{\nu} J_1(2\pi\nu)$,

where $\nu = \sqrt{(u^2 + v^2)}$ represents radial frequency in the 2D (u, v) plane, and J_1 is the first-order Bessel function of the first kind.



(a)

(b)

Figure 2.31: (a) Image of a circular disc. The radius of the disc is 10 pixels; the size of the image is 128×128 pixels. (b) Log-magnitude spectrum of the image in (a). See also Figures 2.32 and 2.33.

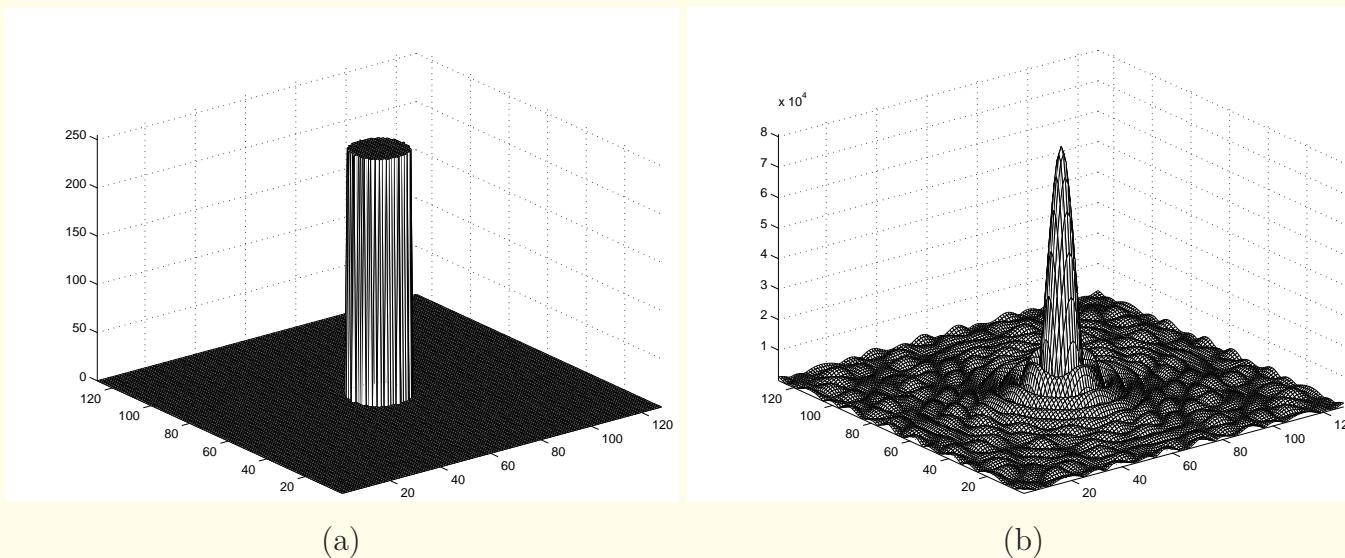
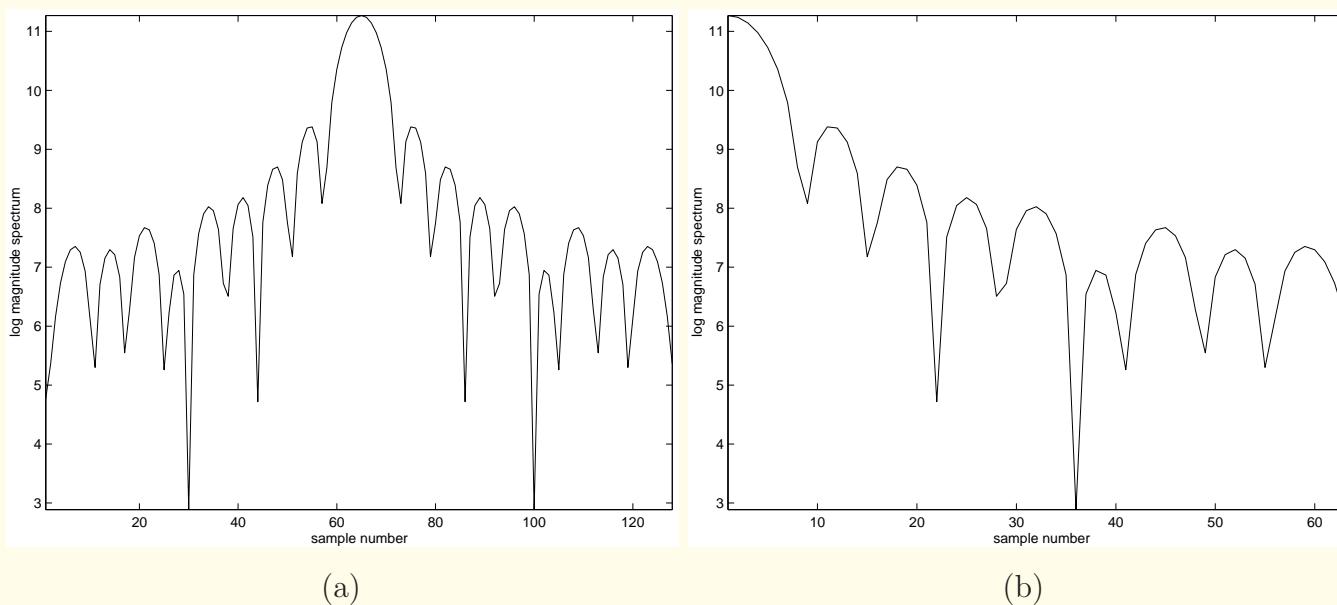


Figure 2.32: (a) Mesh plot of the circular disc in Figure 2.31 (a). The radius of the disc is 10 pixels; the size of the image is 128×128 pixels. (b) Magnitude spectrum of the image in (a).



(a)

(b)

Figure 2.33: (a) Profile of the log-magnitude spectrum in Figure 2.31 (b) along the central horizontal axis. (b) Profile in (a) shown only for positive frequencies. The frequency axis is indicated in samples; the true frequency values depend upon the sampling frequency.



Profiles of 2D system transfer functions:

it is common to show only one half of the profile for positive frequencies.

It is to be assumed that the system possesses axial or rotational symmetry; that is, the system is *isotropic*.

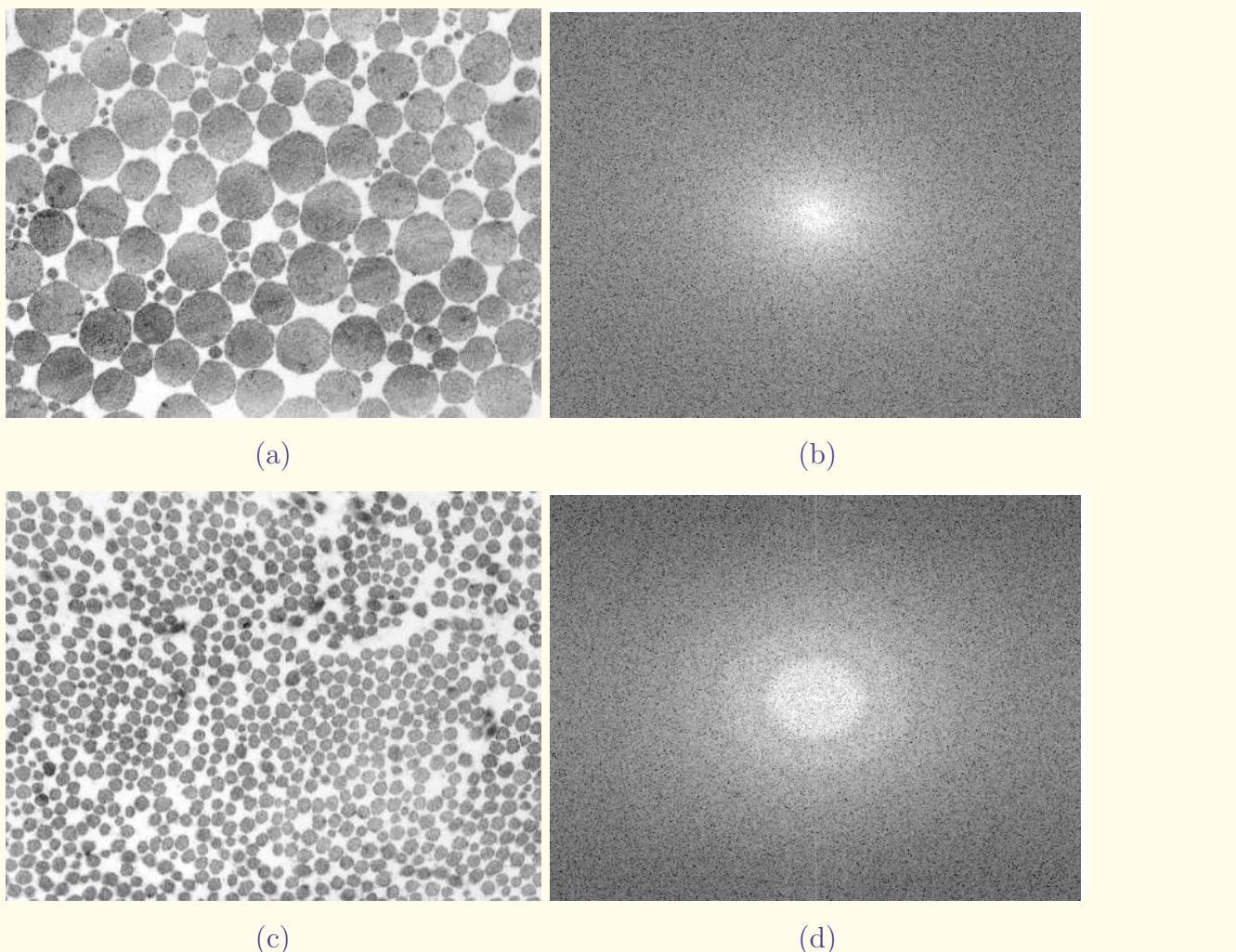
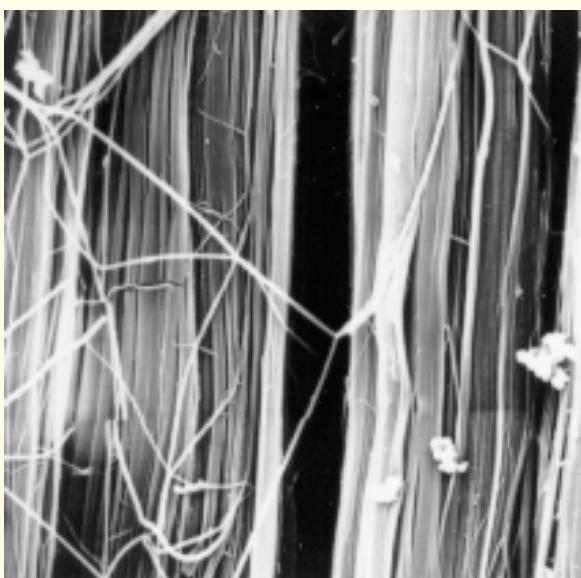
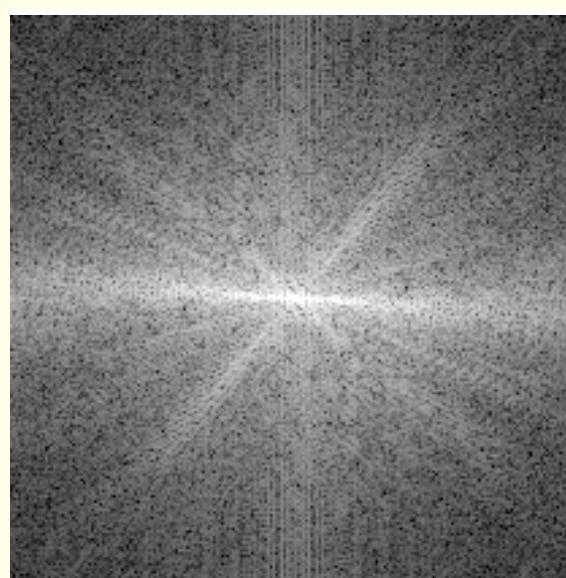


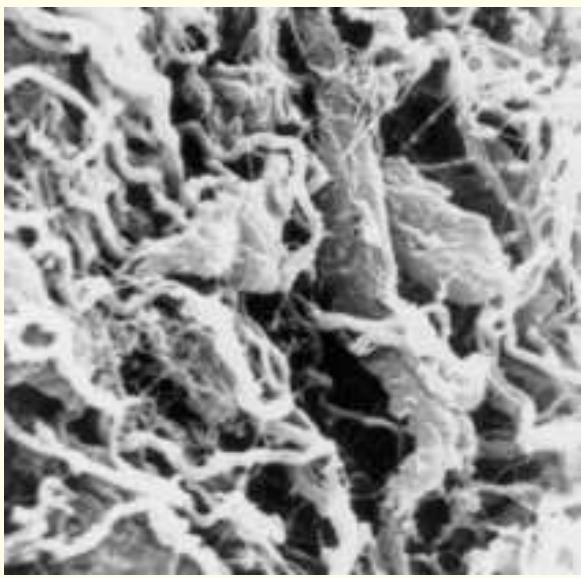
Figure 2.34: (a) TEM image of collagen fibers in a normal rabbit ligament sample. (b) Log-magnitude spectrum of the image in (a). (c) TEM image of collagen fibers in a scar tissue sample. (d) Log-magnitude spectrum of the image in (c). See also Figure 1.5 and Section 1.4.



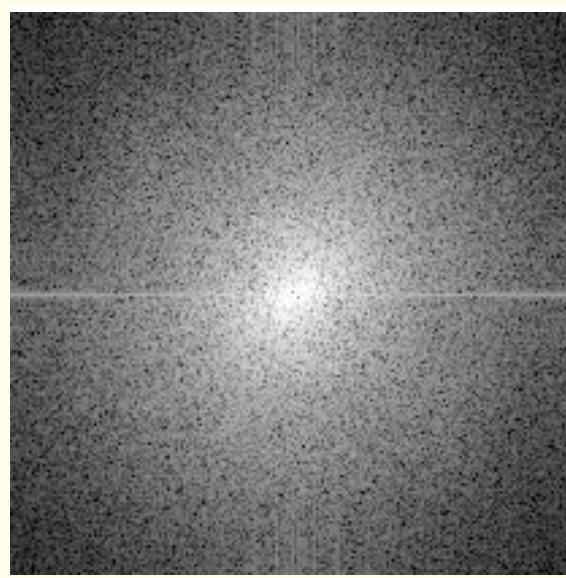
(a)



(b)



(c)



(d)

Figure 2.35: (a) SEM image of collagen fibers in a normal rabbit ligament sample. (b) Log-magnitude spectrum of the image in (a). (c) SEM image of collagen fibers in a scar tissue sample. (d) Log-magnitude spectrum of the image in (c). See also Figure 1.8 and Section 1.4.



2.11.1 *Important properties of the Fourier transform (FT)*

1. The kernel of the FT is separable and symmetric.

Facilitates the evaluation of the 2D DFT as a set of 1D row transforms, followed by a set of 1D column transforms.

$$F(k, l) = \frac{1}{N} \times \quad (2.52)$$

$$\sum_{m=0}^{N-1} \exp\left(-j\frac{2\pi}{N}mk\right) \sum_{n=0}^{N-1} f(m, n) \exp\left(-j\frac{2\pi}{N}nl\right).$$



1D FFT routines may be used to obtain 2D/ MD FT:

$$F(m, l) = N \left[\frac{1}{N} \sum_{n=0}^{N-1} f(m, n) \exp \left(-j \frac{2\pi}{N} nl \right) \right], \quad (2.53)$$

$$F(k, l) = \frac{1}{N} \sum_{m=0}^{N-1} F(m, l) \exp \left(-j \frac{2\pi}{N} mk \right). \quad (2.54)$$

Check if $\frac{1}{N}$ is included in the forward or inverse 1D FFT.



2. Parseval's theorem:

The Fourier transform is an energy-conserving transform.

$$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} |f(x, y)|^2 dx dy = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{\infty} |F(u, v)|^2 du dv. \quad (2.55)$$



3. The inverse Fourier transform (IFT) may be performed using the same FFT routine by taking the forward Fourier transform of the complex conjugate of the given function, and then taking the complex conjugate of the result:

$$f = IFT(F) = [FT\{F^*\}]^*$$



4. The Fourier transform is a linear transform.

Images are often corrupted by additive noise:

$$g(x, y) = f(x, y) + \eta(x, y). \quad (2.56)$$

$$G(u, v) = F(u, v) + \eta(u, v). \quad (2.57)$$



Most real-life images have a large portion of their energy concentrated at $(u, v) = (0, 0)$ in a low-frequency region.

Edges, sharp features, and small-scale or fine details lead to increased strength of high-frequency components.

Random noise has a spectrum that is equally spread all over the frequency space (flat, uniform, or “white” spectrum).

Indiscriminate removal of high-frequency components could cause blurring of edges and the loss of the fine details.



5. The DFT and its inverse are periodic signals:

$$\begin{aligned} F(k, l) &= F(k \pm \alpha N, l) = F(k, l \pm \alpha N) \\ &= F(k \pm \alpha N, l \pm \beta N), \end{aligned} \quad (2.58)$$

where α and β are integers.



6. The Fourier transform is conjugate-symmetric for images with real values:

$$F(-k, -l) = F^*(k, l). \quad (2.59)$$

$$|F(-k, -l)| = |F(k, l)|,$$

$$\angle F(-k, -l) = -\angle F(k, l).$$

The magnitude spectrum is even symmetric.

The phase spectrum is odd symmetric.



7. A spatial shift or translation applied to an image leads to an additional linear phase component in its Fourier transform.

The magnitude spectrum is unaffected.

If $f(m, n) \Leftrightarrow F(k, l)$ are a Fourier-transform pair, we have

$$f(m - m_o, n - n_o) \Leftrightarrow F(k, l) \exp \left[-j \frac{2\pi}{N} (km_o + ln_o) \right], \quad (2.60)$$

where (m_o, n_o) is the shift applied in the space domain.



Conversely, we also have

$$f(m, n) \exp \left[j \frac{2\pi}{N} (k_o m + l_o n) \right] \Leftrightarrow F(k - k_o, l - l_o). \quad (2.61)$$

This property has important implications in the modulation of 1D signals for transmission and communication; it does not have a similar application with 2D images.



8. $F(0, 0)$ gives the average value of the image;
a scale factor may be required depending upon the definition
of the DFT used.



9. For display purposes, $\log_{10}[1 + |F(k, l)|^2]$ is often used; the addition of unity (to avoid taking the log of zero), and the squaring may sometimes be dropped.
- It is also common to fold or shift the spectrum to bring the $(0, 0)$ frequency point (the “DC” point) to the center. Folding of the spectrum could be achieved by multiplying the image $f(m, n)$ with $(-1)^{(m+n)}$ before the FFT is computed. Because the indices m and n are integers, this amounts to merely changing the signs of alternate pixels.



This is related to the property in Equation 2.61 with

$$k_o = l_o = N/2:$$

$$\exp\left[j\frac{2\pi}{N}(k_o m + l_o n)\right] = \exp[j\pi(m+n)] = (-1)^{(m+n)}, \quad (2.62)$$

$$f(m, n) (-1)^{(m+n)} \Leftrightarrow F(k - N/2, l - N/2). \quad (2.63)$$



10. Rotation of an image leads to a corresponding rotation of the Fourier spectrum.

$$f(m_1, n_1) \Leftrightarrow F(k_1, l_1), \quad (2.64)$$

$$m_1 = m \cos \theta + n \sin \theta; \quad n_1 = -m \sin \theta + n \cos \theta; \quad (2.65)$$

$$k_1 = k \cos \theta + l \sin \theta; \quad l_1 = -k \sin \theta + l \cos \theta. \quad (2.66)$$



11. Scaling an image leads to an inverse scaling of its Fourier transform:

$$f(am, bn) \Leftrightarrow \frac{1}{|ab|} F\left(\frac{k}{a}, \frac{l}{b}\right), \quad (2.67)$$

where a and b are scalar scaling factors.

The shrinking of an image leads to an expansion of its spectrum, with increased high-frequency content.

On the contrary, if an image is enlarged, its spectrum is shrunk, with reduced high-frequency energy.



12. Linear shift-invariant systems and convolution:

Most imaging systems may be modeled as linear and shift-invariant or position-invariant systems that are completely characterized by their PSFs.

The output of such a system is given as the convolution of the input image with the PSF:

$$g(m, n) = h(m, n) * f(m, n) \quad (2.68)$$

$$= \sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} h(\alpha, \beta) f(m - \alpha, n - \beta).$$



Upon Fourier transformation, the convolution maps to the multiplication of the two spectra:

$$G(k, l) = H(k, l) F(k, l). \quad (2.69)$$

$$h(x, y) * f(x, y) \Leftrightarrow H(u, v) F(u, v), \quad (2.70)$$

expressed now in the continuous coordinates (x, y) and (u, v) .

The convolution \Leftrightarrow multiplication property with the DFT implies periodic or circular convolution;

Circular convolution may be made to be equivalent to linear convolution by zero-padding.



13. Multiplication of images in the space domain is equivalent to the convolution of their Fourier transforms:

$$f_1(x, y) f_2(x, y) \Leftrightarrow F_1(u, v) * F_2(u, v). \quad (2.71)$$

Some types of noise get multiplied with the image.

When a transparency, such as an X-ray image on film, is viewed using a light box, the resulting image $g(x, y)$ may be modeled as the product of the transparency or transmittance function $f(x, y)$ with the light source intensity field $s(x, y)$:

$$g(x, y) = f(x, y) s(x, y).$$



If $s(x, y)$ is absolutely uniform with a value A , its Fourier transform will be an impulse:

$$S(u, v) = A \delta(u, v).$$

The convolution of $F(u, v)$ with $A \delta(u, v)$ will have no effect on the spectrum except scaling by the constant A .

If the source is not uniform, the viewed image will be a distorted version of the original:

$$G(u, v) = F(u, v) * S(u, v).$$



14. The correlation of two images $f(m, n)$ and $g(m, n)$ is

$$\gamma_{f,g}(\alpha, \beta) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) g(m+\alpha, n+\beta). \quad (2.72)$$

Correlation is useful in the comparison of images where features that are common to the images may be present with a spatial shift (α, β) .

$$\Gamma_{f,g}(k, l) = F(k, l) G^*(k, l). \quad (2.73)$$



A related measure, known as the *correlation coefficient* and useful in *template matching* and image classification, is

$$\gamma = \frac{\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) g(m, n)}{\left[\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f^2(m, n) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} g^2(m, n) \right]^{\frac{1}{2}}}. \quad (2.74)$$

Here, it is assumed that the two images f and g are aligned and registered, and are of the same scale and orientation.



15. Differentiation of an image results in the extraction of edges and highpass filtering:

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &\Leftrightarrow j2\pi u F(u, v); \\ \frac{\partial f(x, y)}{\partial y} &\Leftrightarrow j2\pi v F(u, v).\end{aligned}\tag{2.75}$$

Gain of the filter increases linearly with frequency u or v .



Derivatives are approximated by differences:

$$\begin{aligned}f'_y(m, n) &\approx f(m, n) - f(m - 1, n), \\f'_x(m, n) &\approx f(m, n) - f(m, n - 1),\end{aligned}\quad (2.76)$$

(using matrix notation).



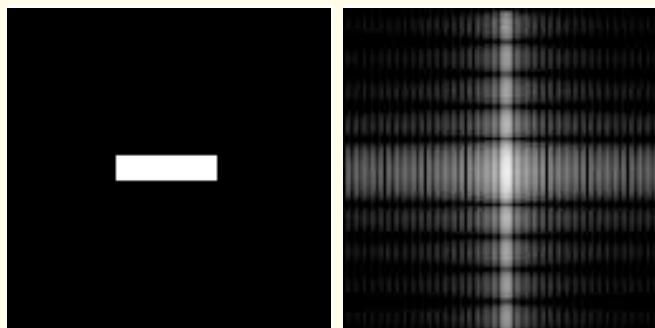
Operators based upon differences could cause negative pixel values in the result.

In order to display the result as an image, it will be necessary to map the full range of the pixel values, including the negative values, to the display range available.

The magnitude of the result may also be displayed if the sign of the result is not important.

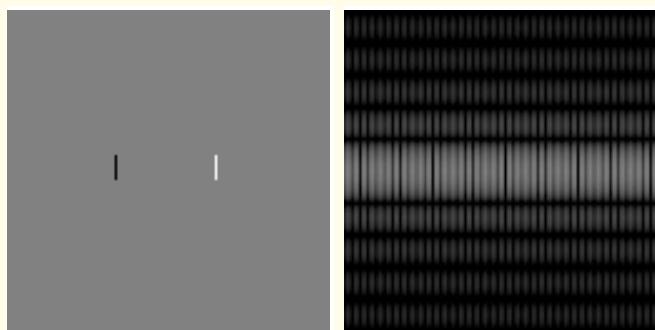
Differentiation results in the removal of the intensity information from the image.

The values of the spectrum for $u = 0$ or $v = 0$ are set to zero.



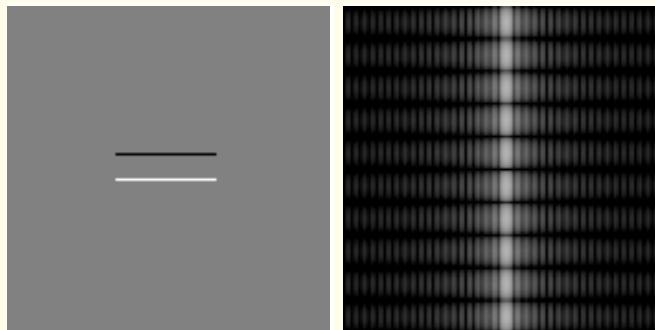
(a)

(b)



(c)

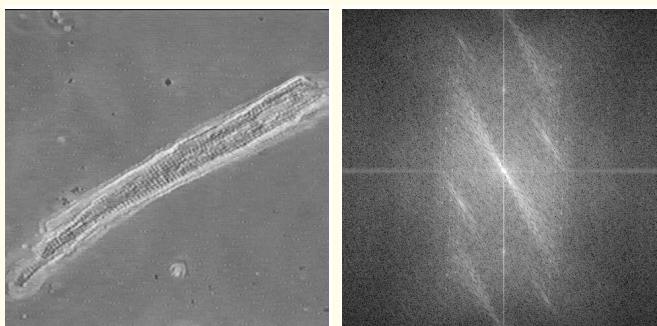
(d)



(e)

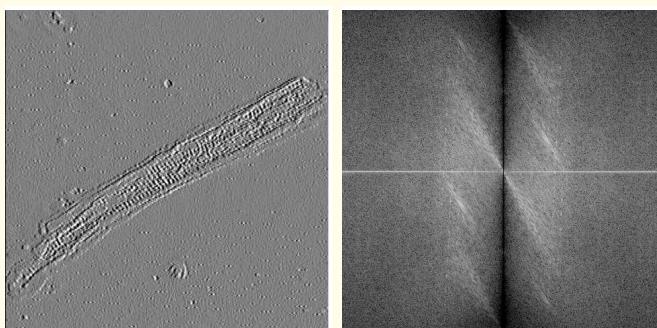
(f)

Figure 2.36: (a) Image of a rectangular box. (c) Horizontal and (e) vertical derivatives of the image in (a), respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. The images in (c) and (e) were obtained by mapping the range $[-200, 200]$ to the display range of $[0, 255]$. Negative differences appear in black, positive differences in white. The spectra show values in the range $[5, 12]$ mapped to $[0, 255]$.



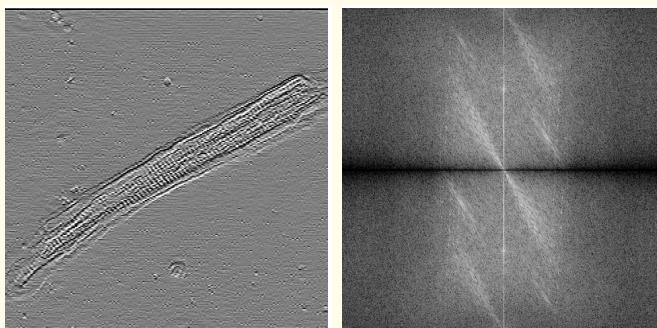
(a)

(b)



(c)

(d)



(e)

(f)

Figure 2.37: (a) Image of a myocyte. (c) Horizontal and (e) vertical derivatives of the image in (a), respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. Images in (c) and (e) were obtained by mapping the range $[-20, 20]$ to the display range of $[0, 255]$. The spectra show values in the range $[3, 12]$ mapped to $[0, 255]$.

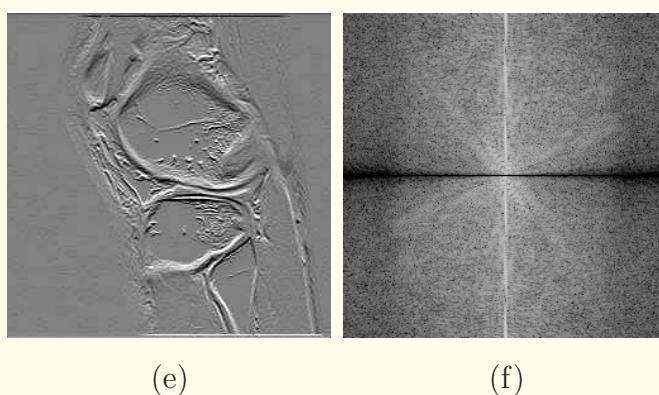
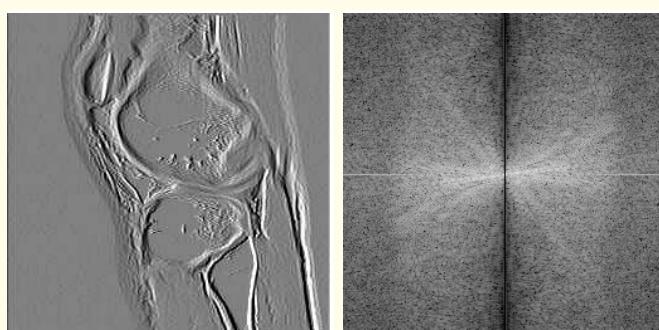
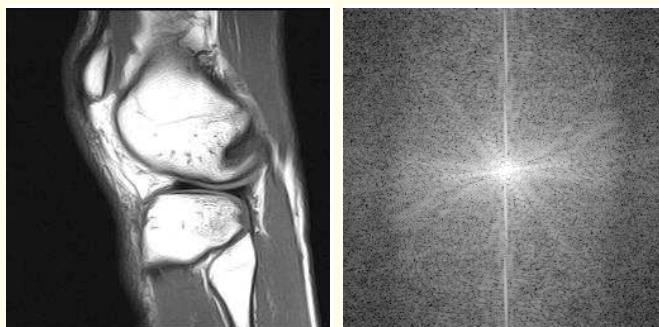


Figure 2.38: (a) MR image of a knee. (c) Horizontal and (e) vertical derivatives of the image in (a), respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. The images in (c) and (e) were obtained by mapping the range $[-50, 50]$ to the display range of $[0, 255]$. Negative differences appear in black, positive differences in white. The spectra show values in the range $[3, 12]$ mapped to $[0, 255]$.



16. The *Laplacian* of an image:

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \quad (2.77)$$

$$\nabla^2 f(x, y) \Leftrightarrow -(2\pi)^2(u^2 + v^2)F(u, v). \quad (2.78)$$

The spectrum of the image is multiplied by $(u^2 + v^2)$, which is isotropic and increases quadratically with frequency.

High-frequency components are amplified.

Omnidirectional operator: detects edges in all directions.



The second derivatives may be approximated as follows:

Taking the derivative of the expression for $f'_y(m, n)$ in Equation 2.76 for the second time, we get

$$f''_y(m, n)$$

$$\begin{aligned} &\approx f(m, n) - f(m - 1, n) - [f(m - 1, n) - f(m - 2, n)] \\ &= f(m, n) - 2f(m - 1, n) + f(m - 2, n) \end{aligned} \quad (2.79)$$

(using matrix notation).

Causality is usually not of concern in image processing: desirable to have operators use collections of pixels that are centered about the pixel being processed.



Applying a shift of one pixel to the result above, we get

$$f_y''(m, n) \quad (2.80)$$

$$\approx f(m + 1, n) - 2 f(m, n) + f(m - 1, n)$$

$$= f(m - 1, n) - 2 f(m, n) + f(m + 1, n).$$

$$f_x''(m, n) \approx f(m, n - 1) - 2 f(m, n) + f(m, n + 1). \quad (2.81)$$



The Laplacian could then be implemented as

$$\begin{aligned} f_L(m, n) = & f(m - 1, n) + f(m, n - 1) - 4f(m, n) \\ & + f(m + 1, n) + f(m, n + 1). \end{aligned} \quad (2.82)$$

\equiv convolving the image with the 3×3 mask or operator

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (2.83)$$

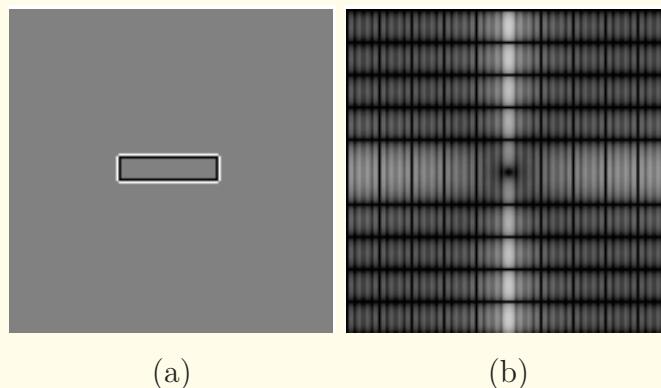
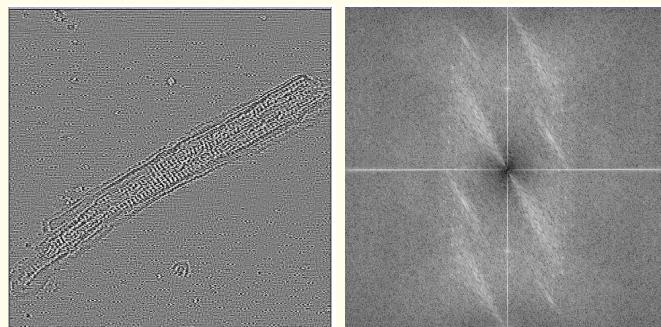


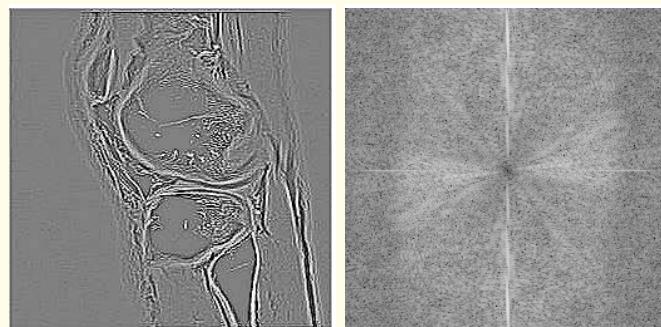
Figure 2.39: (a) Laplacian of the rectangle image in Figure 2.36 (a). (b) Log-magnitude spectrum of the image in (a).



(a)

(b)

Figure 2.40: (a) Laplacian of the myocyte image in Figure 2.37 (a). (b) Log-magnitude spectrum of the image in (a).



(a)

(b)

Figure 2.41: (a) Laplacian of the MR image in Figure 2.38 (a). (b) Log-magnitude spectrum of the image in (a).



17. Integration of an image leads to smoothing or blurring, and lowpass filtering:

$$\int_{\alpha=-\infty}^x f(\alpha, y) d\alpha \Leftrightarrow \frac{1}{j2\pi u} F(u, v), \quad (2.84)$$

$$\int_{\beta=-\infty}^y f(x, \beta) d\beta \Leftrightarrow \frac{1}{j2\pi v} F(u, v). \quad (2.85)$$

The weighting factors that apply to $F(u, v)$ diminish with increasing frequency, and hence high-frequency components are attenuated by this operation.



The integration of an image from $-\infty$ to the current x or y position is seldom encountered in practice.

Instead, it is common to encounter the integration of an image over a small region or aperture surrounding the current position, in the form

$$g(x, y) = \frac{1}{AB} \int_{\alpha=-A/2}^{A/2} \int_{\beta=-B/2}^{B/2} f(x + \alpha, y + \beta) d\alpha d\beta, \quad (2.86)$$

where the region of integration is a rectangle of size $A \times B$.

The normalization factor $\frac{1}{AB}$ leads to the average intensity being computed over the area of integration.

This may be interpreted as a moving-average (MA) filter.



Averaging over a 3×3 aperture or neighborhood:

$$g(m, n) = \frac{1}{9} \sum_{\alpha=-1}^1 \sum_{\beta=-1}^1 f(m + \alpha, n + \beta). \quad (2.87)$$

$$\begin{aligned} g(m, n) = & \frac{1}{9} \times \\ & [f(m - 1, n - 1) + f(m - 1, n) + f(m - 1, n + 1) \\ & + f(m, n - 1) + f(m, n) + f(m, n + 1) \\ & + f(m + 1, n - 1) + f(m + 1, n) + f(m + 1, n + 1)]. \end{aligned} \quad (2.88)$$



Equivalent to convolution of the image $f(m, n)$ with

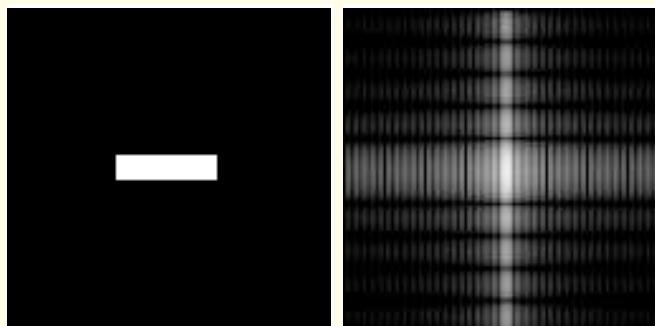
$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad (2.89)$$

which may be viewed as the PSF of a filter.

Equivalent to multiplication of the Fourier transform of the image with a 2D sinc function.

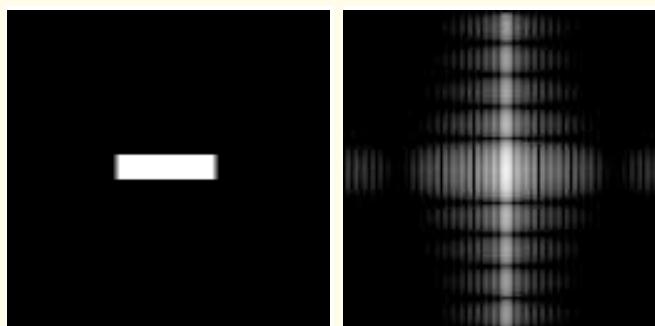


Integration or averaging only along the horizontal or vertical directions may be performed via convolution with the arrays $\frac{1}{3} [1, 1, 1]$ or $\frac{1}{3} [1, 1, 1]^T$.



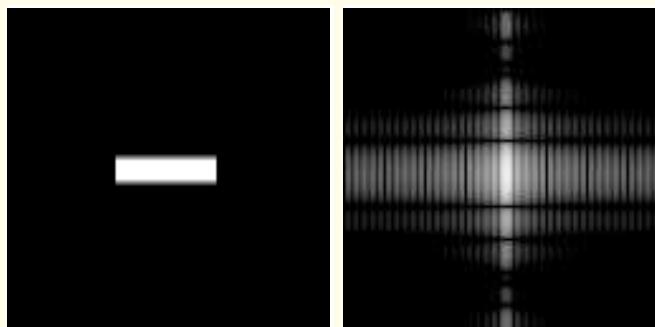
(a)

(b)



(c)

(d)



(e)

(f)

Figure 2.42: (a) Image of a rectangular box. Results of averaging using three pixels in the (c) horizontal and (e) vertical directions, respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. The spectra show values in the range [5, 12] mapped to [0, 255].

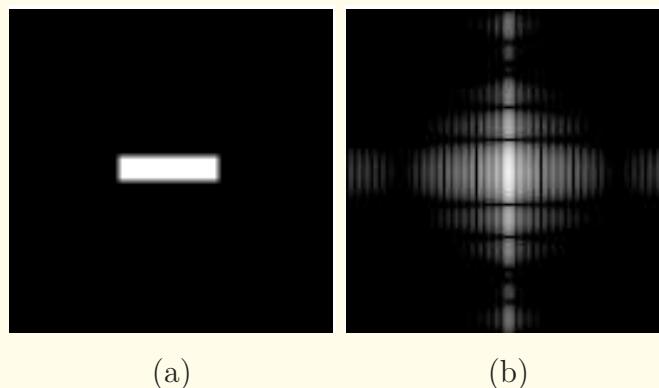


Figure 2.43: (a) Result of 3×3 averaging of the rectangle image in Figure 2.42 (a). (b) Log-magnitude spectrum of the image in (a).

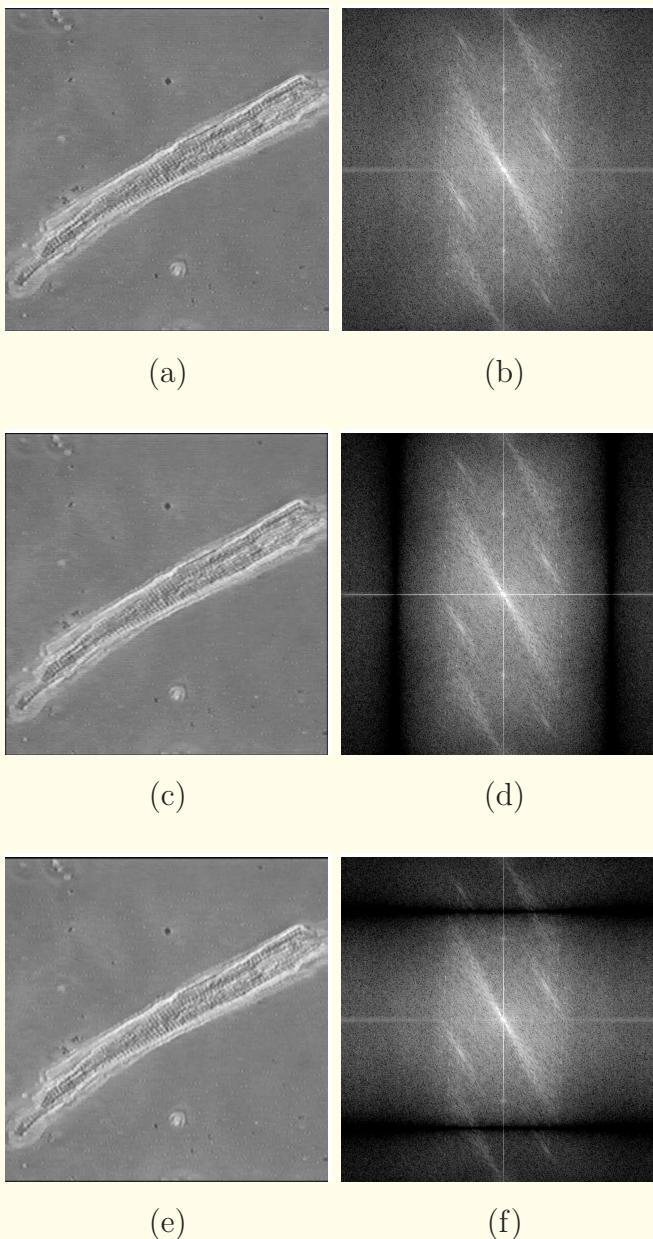
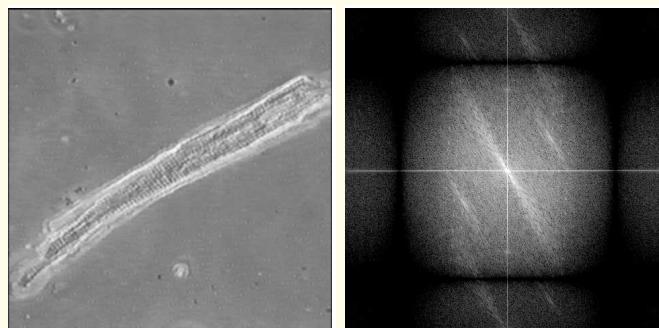


Figure 2.44: (a) Image of a myocyte. Results of averaging using three pixels in the (c) horizontal and (e) vertical directions, respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. The spectra show values in the range [3, 12] mapped to [0, 255].



(a)

(b)

Figure 2.45: (a) Result of 3×3 averaging of the myocyte image in Figure 2.44 (a). (b) Log-magnitude spectrum of the image in (a).

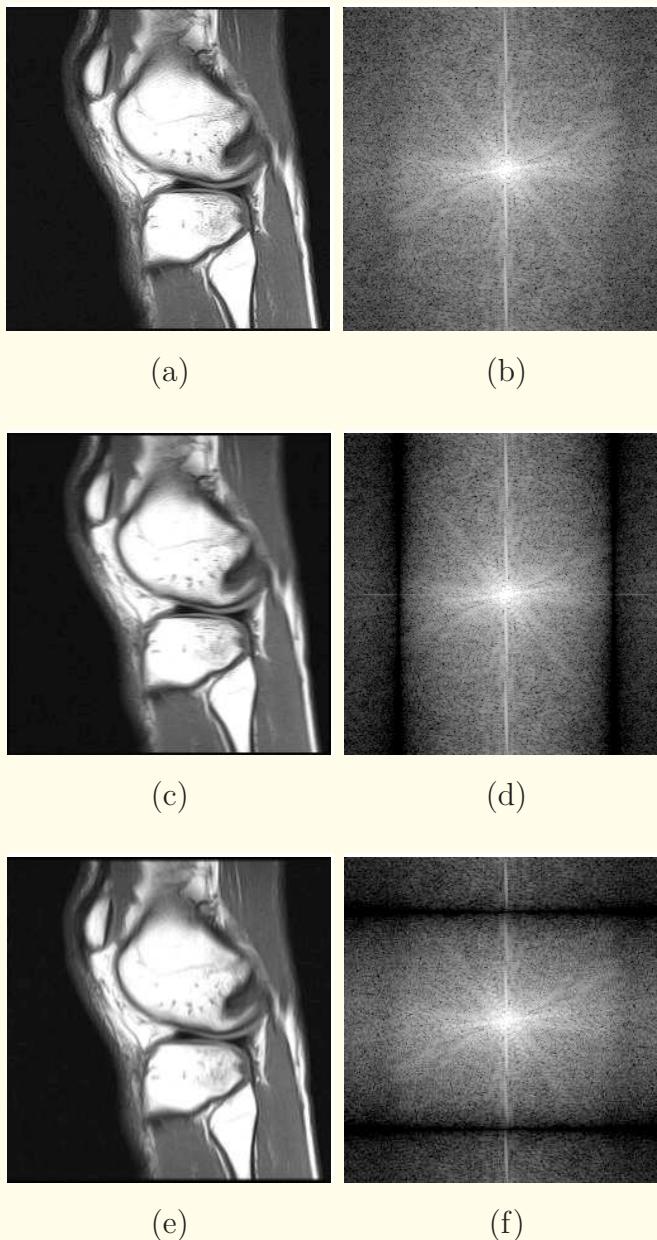
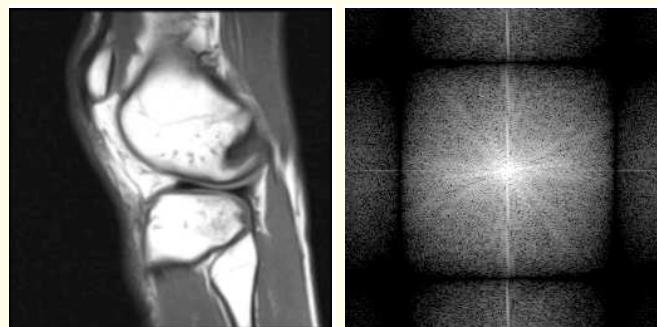


Figure 2.46: (a) MR image of a knee. Results of averaging using three pixels in the (c) horizontal and (e) vertical directions, respectively. (b), (d), and (f): Log-magnitude spectra of the images in (a), (c), and (e), respectively. The spectra show values in the range [3, 12] mapped to [0, 255].



(a)

(b)

Figure 2.47: (a) Result of 3×3 averaging of the knee MR image in Figure 2.46 (a). (b) Log-magnitude spectrum of the image in (a).



2.12 Modulation Transfer Function (MTF)

Analysis of the characteristics of imaging systems, treated as 2D LSI systems, is easier in the frequency or Fourier domain.

$$G(u, v) = H(u, v) F(u, v), \quad (2.90)$$

where $F(u, v)$, $G(u, v)$, and $H(u, v)$ are the 2D Fourier transforms of $f(x, y)$, $g(x, y)$, and $h(x, y)$, respectively.

$H(u, v)$ is known as the *optical transfer function (OTF)*.

OTF is, in general, a complex quantity; its magnitude is the *modulation transfer function (MTF)*.



The widths of a PSF and the corresponding MTF bear an inverse relationship:

the greater the blur, the wider the PSF, the narrower the MTF;

more high-frequency components are attenuated significantly.

Resolution expressed in the frequency domain as a point along the frequency axis beyond which the attenuation is significant.

A larger area under the (normalized) MTF indicates a system with better resolution (more high-frequency components preserved) than a system with a smaller area under the MTF.

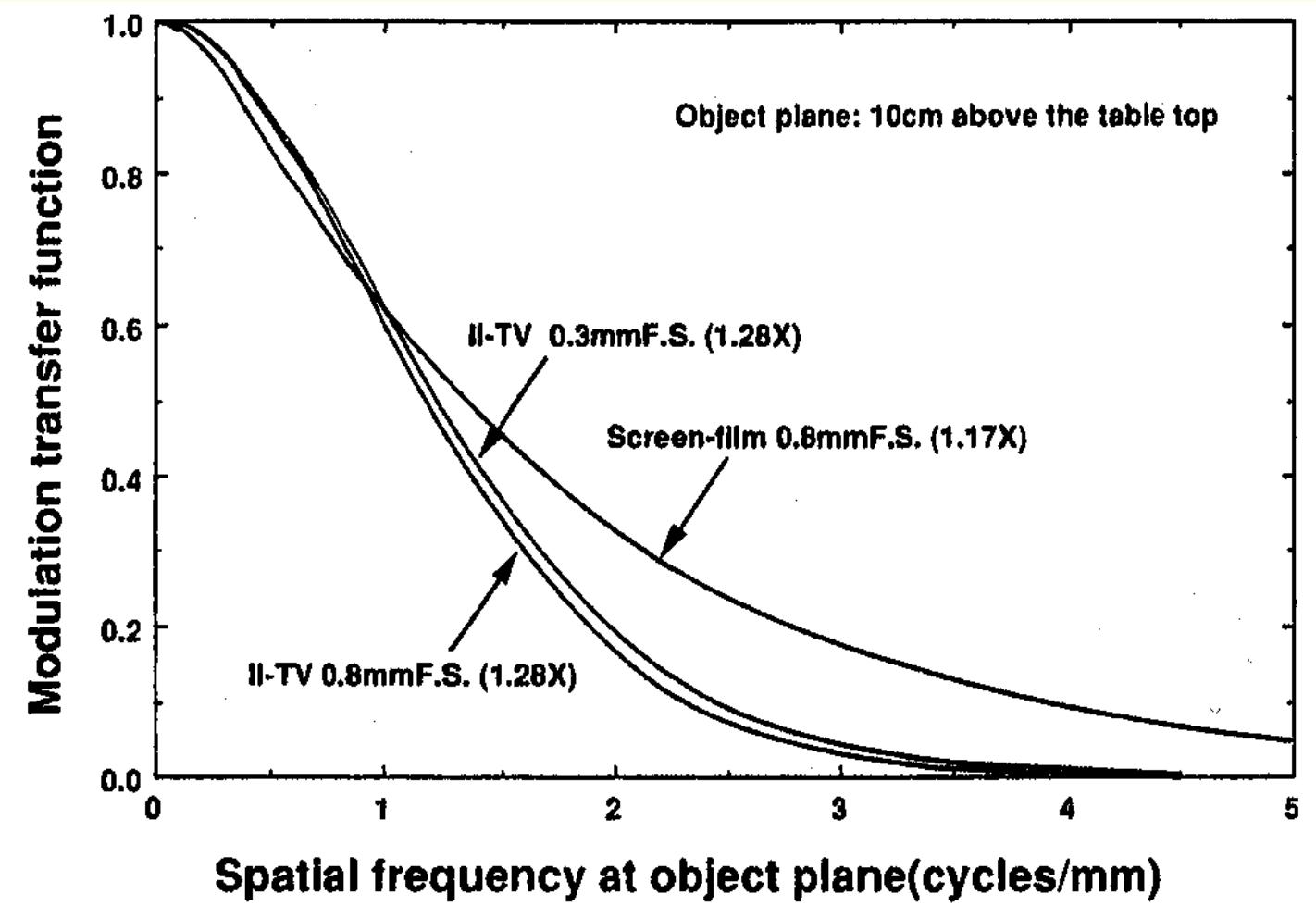


Figure 2.48: MTFs of a DR system (II-TV = image-intensifier television) and a screen-film system at the same X-ray dose. FS = focal spot. Reproduced with permission from Y. Higashida, Y. Baba, M. Hatemura, A. Yoshida, T. Takada, and M. Takahashi, “Physical and clinical evaluation of a $2,048 \times 2,048$ -matrix image intensifier TV digital imaging system in bone radiography”, *Academic Radiology*, 3(10):842–848. 1996. © Association of University Radiologists.



Contrast-detail curve: Based upon experiments with images of square objects of varying thickness.

The threshold object thickness was determined as that related to the lowest-contrast image where radiologists could visually detect the objects in images with a 50% confidence level.

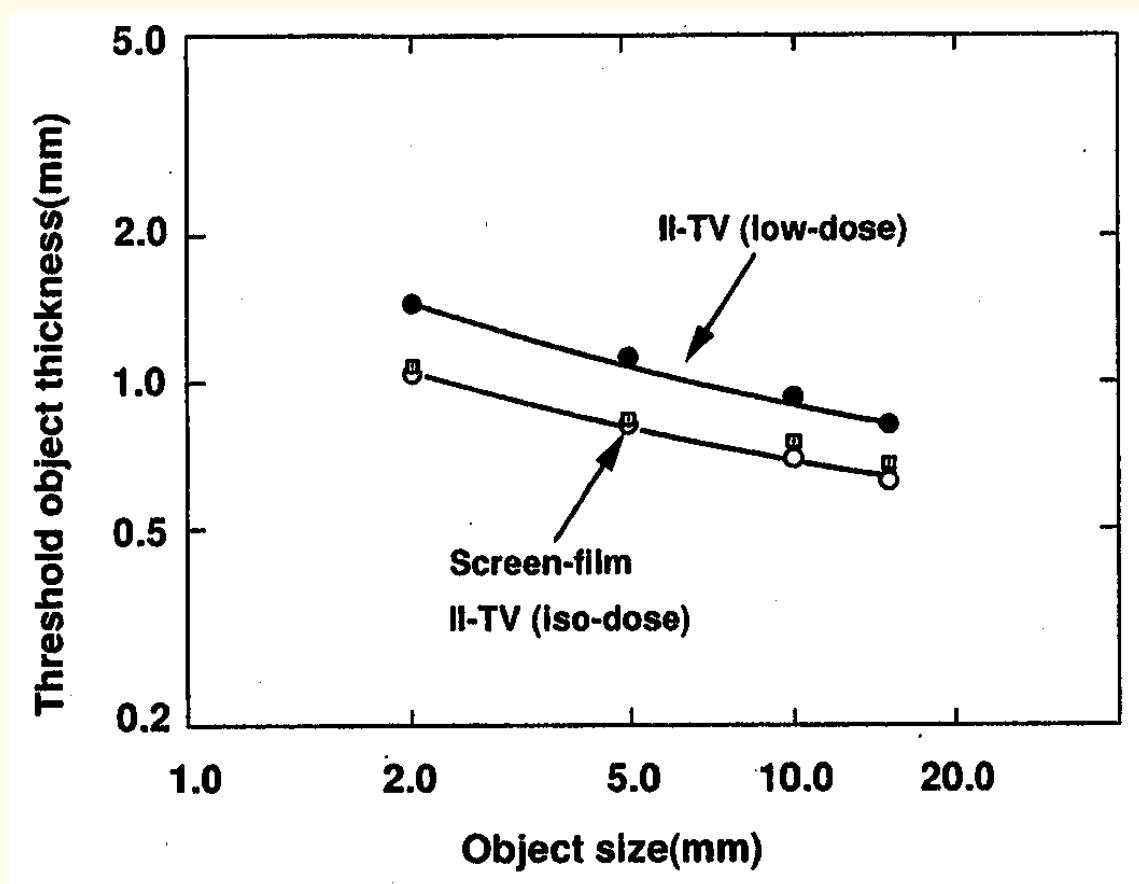


Figure 2.49: Contrast-detail curves of a DR system (II-TV = image-intensifier television) and a screen-film system. The DR system was operated at the same X-ray dose as the screen-film system (iso-dose) and at a low-dose setting. Reproduced with permission from Y. Higashida, Y. Baba, M. Hatemura, A. Yoshida, T. Takada, and M. Takahashi, "Physical and clinical evaluation of a $2,048 \times 2,048$ -matrix image intensifier TV digital imaging system in bone radiography", *Academic Radiology*, 3(10):842–848. 1996. © Association of University Radiologists.

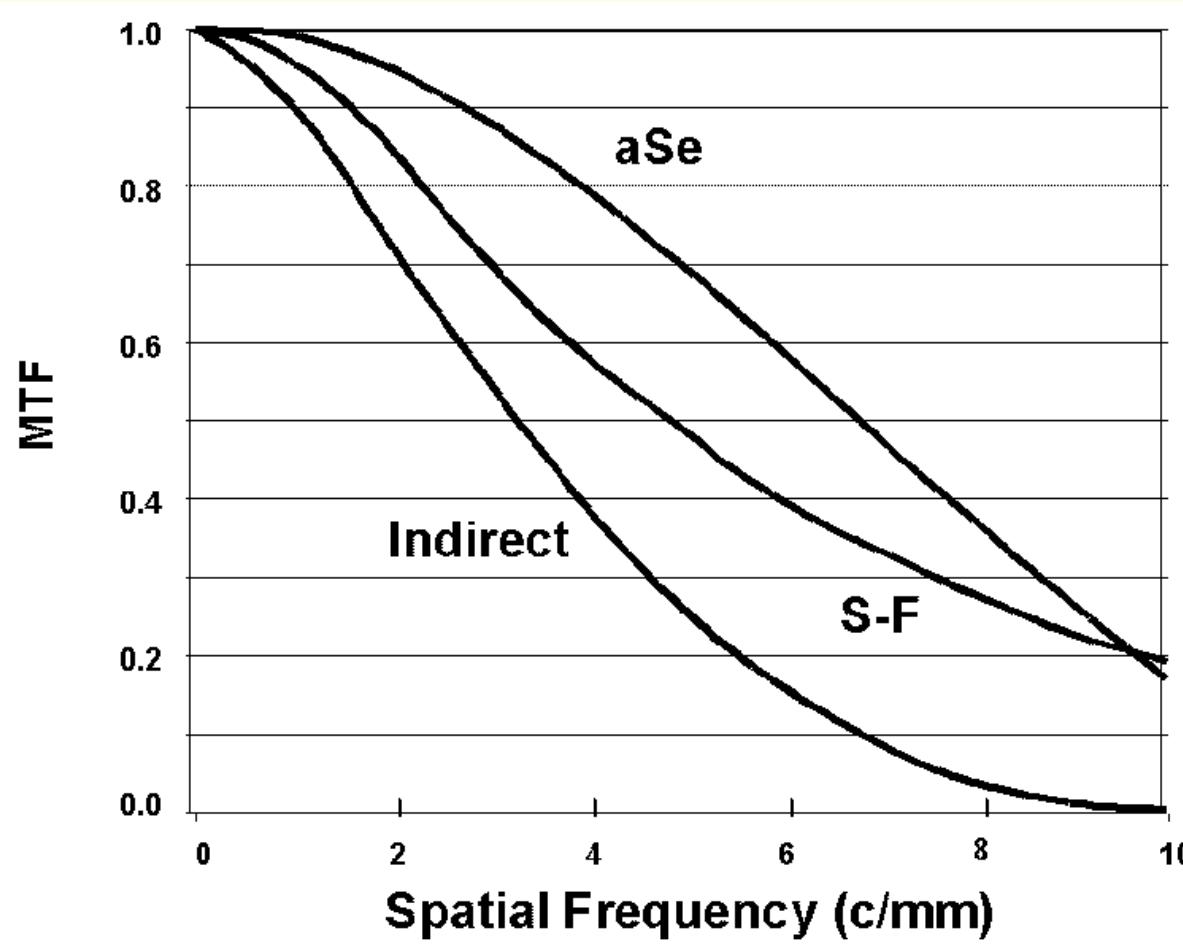


Figure 2.50: MTF curves of an amorphous selenium (aSe) detector system for direct digital mammography, a screen-film system (S-F), and an indirect digital imaging system. $c/mm = \text{cycles}/mm$. Figure courtesy of J.E. Gray, Lorad, Danbury, CT.



MTF of a μ CT system for high-resolution 3D imaging of small samples using synchrotron radiation.

MTF = 0.1 at 55 lp/mm .

Spatial resolution: $(2 \times 55)^{-1} = 0.009\ mm$ or $9\ \mu m$.

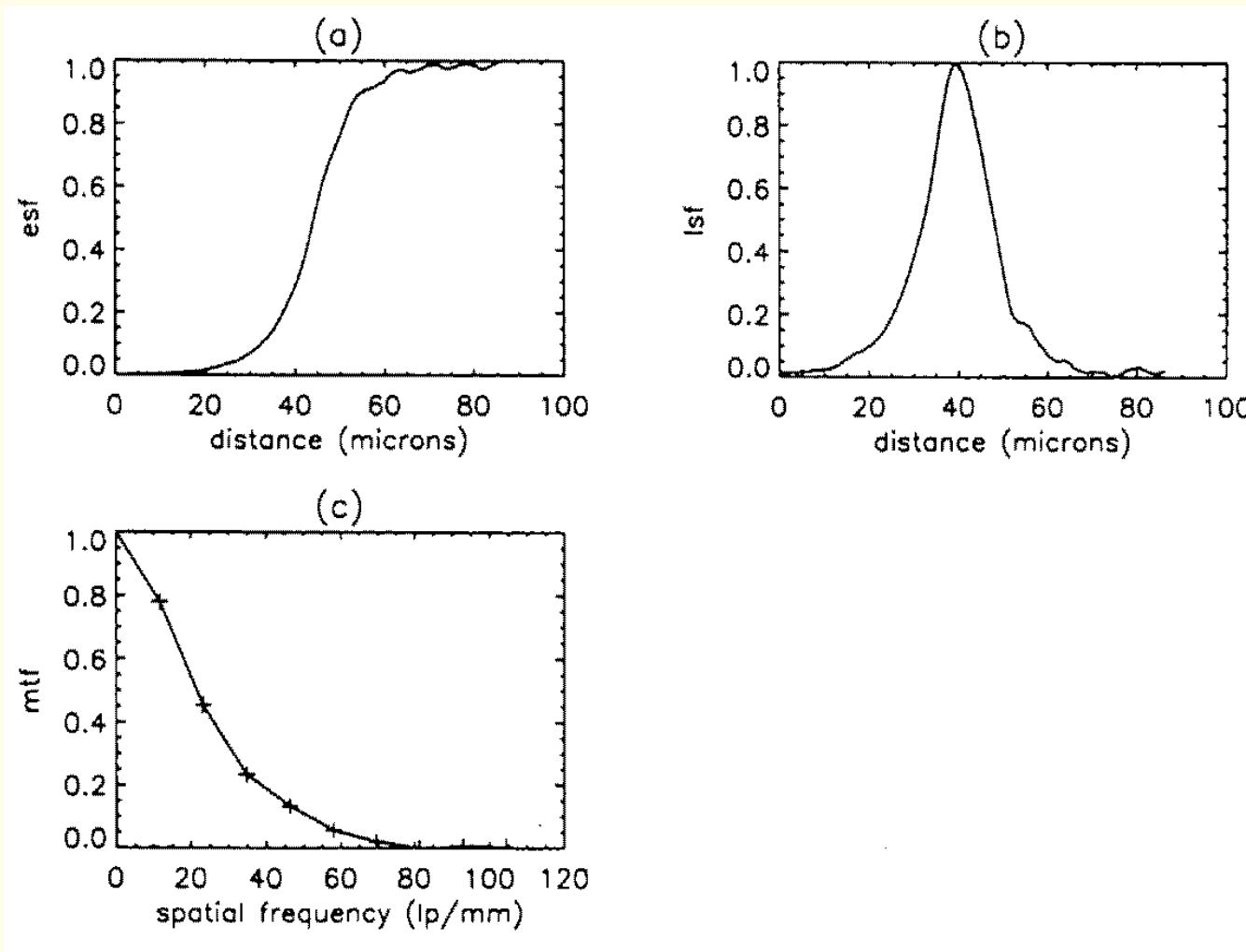


Figure 2.51: (a) Edge spread function, (b) line spread function, and (c) MTF of a μ CT system. 1 micron = 1 μm . Reproduced with permission from M. Pateyron, F. Peyrin, A.M. Laval-Jeantet, P. Spanne, P. Cloetens, and G. Peix, “3D microtomography of cancellous bone samples using synchrotron radiation”, *Proceedings of SPIE 2708: Medical Imaging 1996 – Physics of Medical Imaging*, Newport Beach, CA, pp 417–426. © SPIE.



2.13 Signal-to-Noise Ratio (SNR)

$$g(x, y) = f(x, y) + \eta(x, y). \quad (2.91)$$

Assume noise process is additive and statistically independent of (uncorrelated with) the image process.



Mean:

$$\mu_g = \mu_f + \mu_\eta, \quad (2.92)$$

Usually, the mean of the noise process is zero: $\mu_g = \mu_f$.

Variance:

$$\sigma_g^2 = \sigma_f^2 + \sigma_\eta^2. \quad (2.93)$$



$$SNR_1 = 10 \log_{10} \left[\frac{\sigma_f^2}{\sigma_\eta^2} \right] dB. \quad (2.94)$$

Variance of noise estimated by computing the sample variance of pixels from background areas of the image.

Variance may be computed from the PDF (histogram).



The variance of the image may not provide an appropriate indication of the useful range of variation in the image.

SNR based upon the dynamic range of the image:

$$SNR_2 = 20 \log_{10} \left[\frac{f_{\max} - f_{\min}}{\sigma_\eta} \right] dB. \quad (2.95)$$

Video signals in modern CRT monitors: $\text{SNR} \approx 60 - 70 \text{ dB}$ with noninterlaced frame repetition rate of $70 - 80 \text{ fps}$.



Contrast-to-noise ratio (CNR) is a measure that combines the contrast or the visibility of an object and the SNR:

$$CNR = \frac{\mu_f - \mu_b}{\sigma_b}, \quad (2.96)$$

Simultaneous contrast uses a background that encircles the ROI;

CNR could use a background region located elsewhere.

CNR is well suited to the analysis of X-ray imaging systems:

density of an ROI on a film image depends upon dose;

visibility of an object dependent upon both dose and noise.



2.14 Error-based Measures

Mean-squared error:

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [f(m, n) - g(m, n)]^2. \quad (2.97)$$

Normalized MSE:

$$NMSE = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [f(m, n) - g(m, n)]^2}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [f(m, n)]^2}. \quad (2.98)$$



Normalized error:

$$NE = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |f(m, n) - g(m, n)|}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |f(m, n)|}. \quad (2.99)$$

Laplacian MSE:

$$LMSE = \frac{\sum_{m=1}^{M-2} \sum_{n=1}^{N-2} [f_L(m, n) - g_L(m, n)]^2}{\sum_{m=1}^{M-2} \sum_{n=1}^{N-2} [f_L(m, n)]^2}. \quad (2.100)$$

$f_L(m, n)$ is the Laplacian of $f(m, n)$.



Perceptual MSE:

PMSE defined in a manner similar to LMSE, but with each image replaced with the logarithm of the image convolved with a PSF representing the HVS.

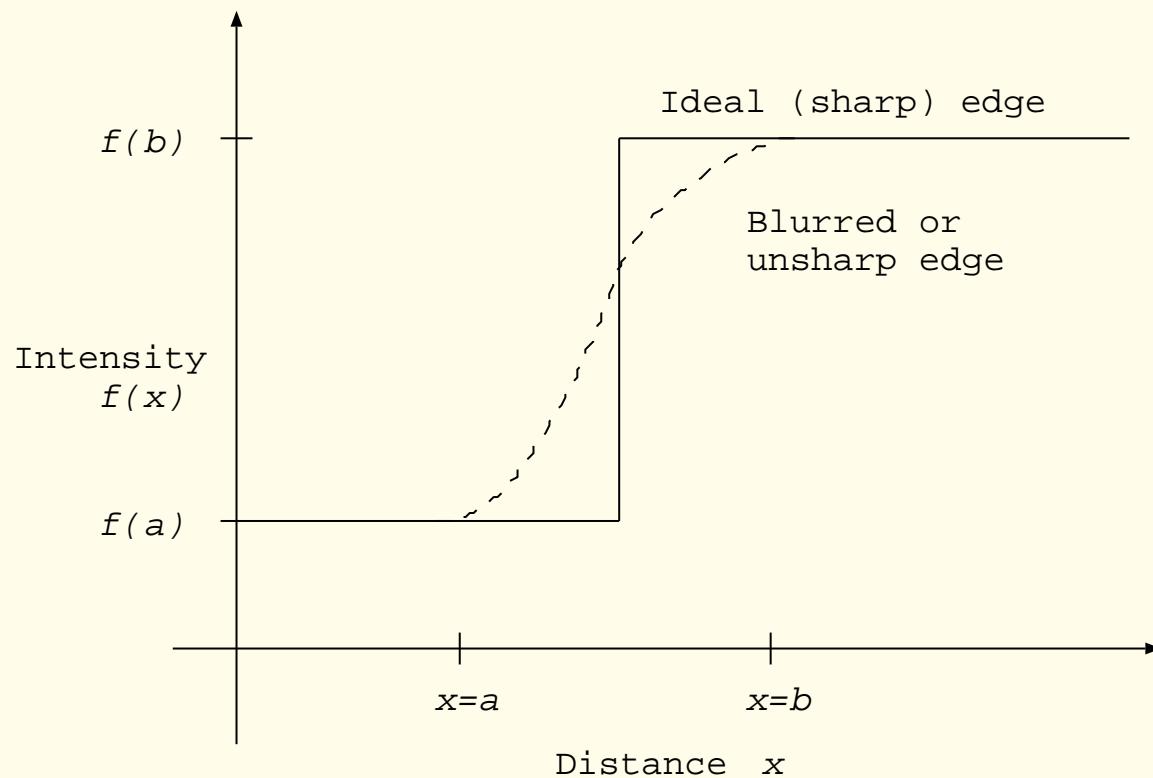
The measures defined above assume the availability of a reference image for comparison in a *before-and-after* manner.



2.15 Application: Image Sharpness and Acutance

Acutance from the edge spread function:

$$A = \frac{1}{f(b) - f(a)} \int_a^b \left[\frac{d}{dx} f(x) \right]^2 dx. \quad (2.101)$$





MTF-based measures of acutance:

MTF of a system composed of LSI systems in series (cascade):

$$H(u, v) = H_1(u, v) \ H_2(u, v) \ \cdots \ H_N(u, v) = \prod_{i=1}^N H_i(u, v). \quad (2.102)$$

$$h(x, y) = h_1(x, y) * h_2(x, y) * \cdots * h_N(x, y). \quad (2.103)$$



System with large gain at high frequencies (sharp output image):

large area under the normalized MTF.

MTF-area-based measures represent the combined effect of all the systems between the image source and the viewer:

independent of the actual image displayed.



Assuming the system to be isotropic, MTF is expressed as a 1D function of radial frequency $\nu = \sqrt{(u^2 + v^2)}$.

$$A_1 = \int_0^{\nu_{\max}} [H_s(\nu) - H_e(\nu)] d\nu, \quad (2.104)$$

where $H_s(\nu)$ is the MTF of the complete chain of systems,

$H_e(\nu)$ is the MTF threshold of the eye,

ν is the radial frequency at the eye of the observer, and

ν_{\max} is given by the condition $H_s(\nu_{\max}) = H_e(\nu_{\max})$.



In order to reduce the weighting on high-frequency components, replace the difference between the MTFs with their ratio:

$$A_2 = \int_0^\infty \frac{H_s(\nu)}{H_e(\nu)} d\nu. \quad (2.105)$$



$$AMTA = 100 + 66 \log_{10} \left[\frac{\int_0^{\infty} H_s(\nu) H_e(\nu) d\nu}{\int_0^{\infty} H_e(\nu) d\nu} \right]. \quad (2.106)$$

MTF of the eye modeled as a Gaussian, $\sigma = 13$ cycles/degree.

AMTA values interpreted as:

- 100 : excellent,
- 90 : good,
- 80 : fair, and
- 70 : just passable.



Barten proposed the evaluation of image quality using the square-root integral:

$$SQRI = \frac{1}{\ln(2)} \int_0^{\nu_{\max}} \left[\frac{H_s(\nu)}{H_e(\nu)} \right]^{\frac{1}{2}} \frac{d\nu}{\nu}. \quad (2.107)$$

ν_{\max} is the maximum frequency to be displayed.



Region-based measure of edge sharpness:

Instead of the traditional difference defined as

$$f'(n) = f(n) - f(n - 1), \quad (2.108)$$

Rangayyan and Elkadiki split the normal at each boundary pixel into a foreground part $f(n)$ and a background part $b(n)$ and defined an averaged gradient as

$$f_d(k) = \frac{1}{N} \sum_{n=1}^N \frac{f(n) - b(n)}{2n}, \quad (2.109)$$

where k is the index of the boundary pixel,
 N is the number of pairs of pixels used along the normal.

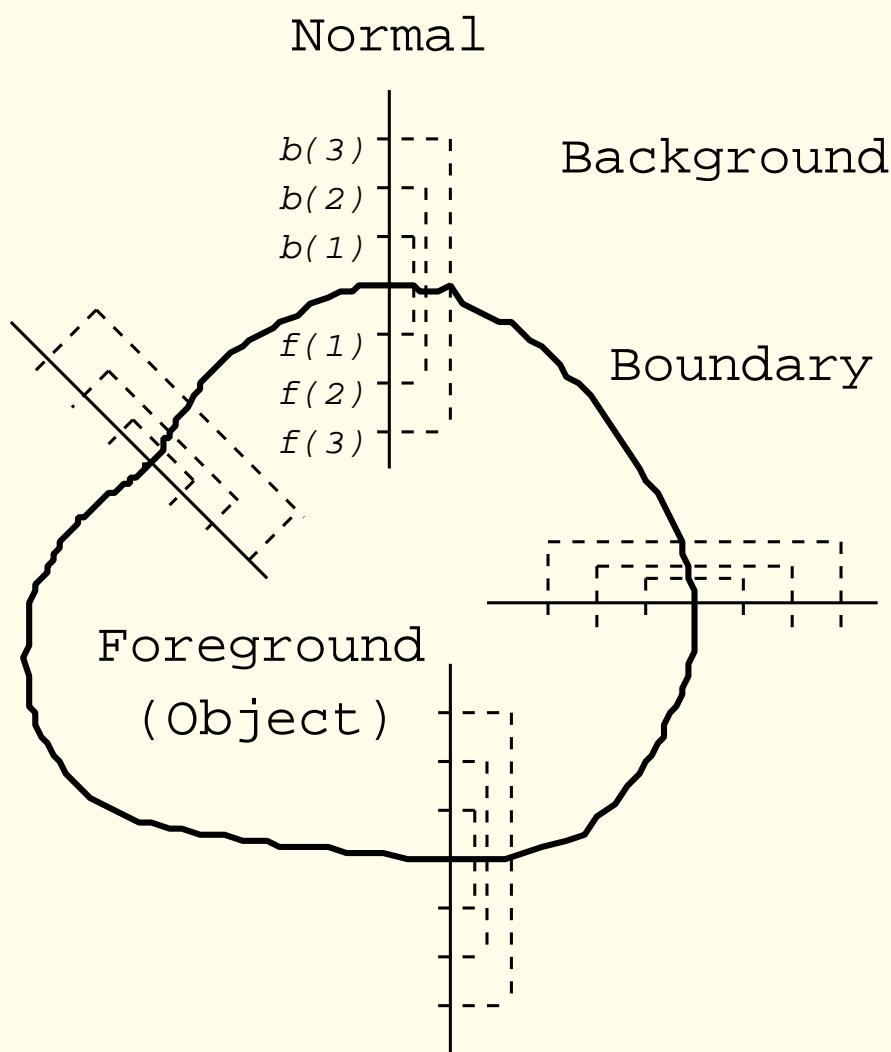


Figure 2.52: Computation of differences along the normals to a region in order to derive a measure of acutance. Four sample normals are illustrated, with three pairs of pixels being used to compute differences along each normal.



The averaged gradient values over all boundary pixels were then combined to obtain a single normalized value of acutance A :

$$A = \frac{1}{d_{\max}} \left[\frac{1}{K} \sum_{k=1}^K f_d^2(k) \right]^{\frac{1}{2}}, \quad (2.110)$$

where K is the number of pixels along the boundary, d_{\max} is the maximum possible gradient value.



Acutance:

- reduced by blurring,
- increased by sharpening,
- not affected significantly by noise,
- correlates well with sharpness as judged by human observers.

