Medical Image Processing Overview

Hongmei Zhu, University of Calgary

Biomedical image processing has experienced dramatic expansion, and has been an interdisciplinary research field attracting expertise from applied mathematics, computer sciences, engineering, statistics, physics, biology and medicine. Computer-aided diagnostic processing has already become an important part of clinical routine. Accompanied by a rush of new development of high technology and use of various imaging modalities, more challenges arise; for example, how to process and analyze a significant volume of images so that high quality information can be produced for disease diagnoses and treatment. The principal objectives of this course are to provide an introduction to basic concepts and techniques for medical image processing and to promote interests for further study and research in medical imaging processing.

1. PRINCIPLES OF MAGNETIC RESONANCE IMAGING

During the past few decades, with the increasing availability of relatively inexpensive computational resources, computed tomography (CT), magnetic resonance imaging (MRI), doppler ultrasound, and various imaging techniques based on nuclear emission (PET (positron emission tomography), SPECT (single photon emission computed tomography), etc) have all been valuable additions to the radiologist's arsenal of imaging tools toward ever more reliable detection and diagnosis of disease. Tomographic imaging principles are rooted in physics, mathematics, computer science, and engineering. Here, we mainly focus on the conceptual overview of the principles of MRI, one of the major imaging modalities currently in use. Excellent detailed discussions can be found in Nishimura [1] and Liang and Lauterbur [2].

What is MRI? Briefly, it is a tomographic imaging technique that produces images of internal physical and chemical characteristics of an object from externally measured nuclear magnetic resonance (NMR) signals. MR imaging is based on the well-known NMR phenomenon independently discovered by Bloch and his coworkers at Stanford [3] and Purcell and his colleagues at Harvard [4] in 1946. Bloch and Purcell shared the Nobel prize in physics for this discovery in 1952.

An MR scanner shown in Fig. 1.1 consists of three main hardware components: a main magnet, a magnetic field gradient system, and a radio-frequency (RF) system. The main magnet is a permanent magnet. It generates a strong uniform static field, referred to as the B_0 field, for polarization of nuclear spins in an object. Most imaging systems operate at fixed field strength in units of Tesla (1 Tesla (T) = 10^4 Gauss (G)). The magnetic field gradient system normally consists of three orthogonal gradient coils, G_x , G_y and G_z, essential for signal localization. The gradient field strength is usually less than 1 G/cm. The RF system consists of a transmitter coil that is capable of generating a rotating magnetic field, referred to as the B_1 field, for exciting a spin system, and a receiver coil that converts a precessing magnetization into an electrical signal. On human imaging systems, the strength of B_1 is typically a small fraction of a Gauss.

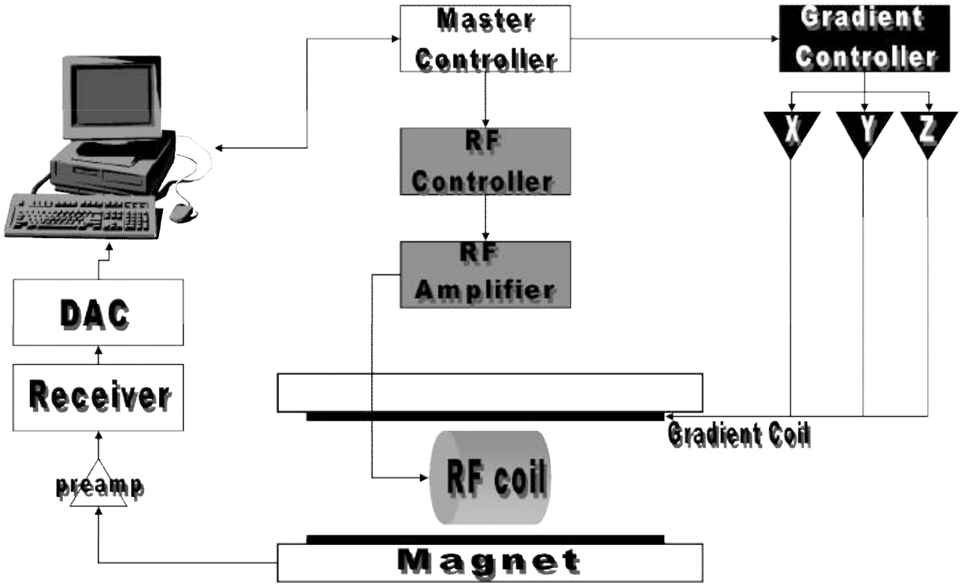

Fig. 1.1 Simplified drawing of the basic instrumentation.

While an MRI system is rather complex, the underlying imaging principles of such a system are rather straightforward, especially from the signal processing point of view. The imaging process involves forward and inverse Fourier transforms as shown in Fig. 1.2.

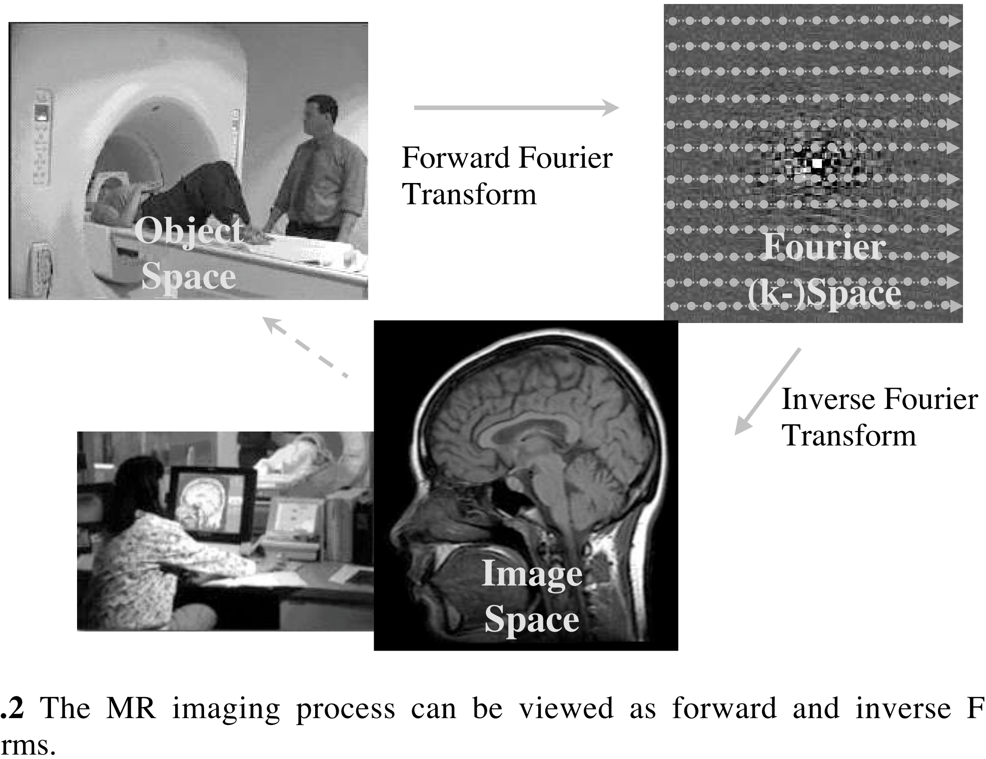

Fig. 1.2 The MR imaging process can be viewed as forward and inverse Fourier transforms.

The imaging process does not involve the use of ionizing radiation and hence does not have the associated potential harmful effects. MR signals used for image formation come directly from the object itself. To understand how MR signals are generated requires understanding of quantum mechanics in the classical vector model.

Atoms with an odd number of protons and/or an odd number of neutrons possess a nuclear spin angular momentum, and therefore exhibit the MR phenomenon. Qualitatively, these nucleons can be visualized as spinning charged spheres that give rise to a small magnetic moment shown in Fig. 3, referred to as spins. In biological specimens, hydrogen (^1H), with a single proton, is the most abundant (the body consists largely of H_2O), the most sensitive and by far the most studied. Therefore, we usually refer to ^1H (proton) imaging.

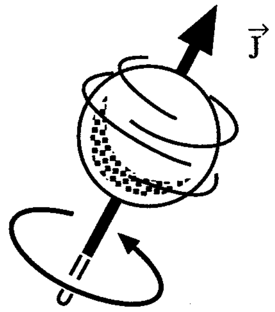

Fig. 1.3 Nuclei with a nonzero spin are regarded as microscopic magnets [2].

In the absence of an external magnetic field, the spins are oriented randomly due to thermal random motion and the net macroscopic magnetic moment is zero (Fig 1.4a). To activate macroscopic magnetism from an object, it is necessary to line up the spin vector. This is accomplished by exposing the object to a strong external magnetic field B_0. The magnetic moment vectors tend to align in the direction of B_0 (referred to as the z-direction) to create a net magnetic moment M (Fig. 1.4b). Also, the spins exhibit resonance at a well-defined frequency, called the Larmor frequency ω. The Larmor frequency is defined as

$$\omega = \gamma B_0 \qquad \text{or} \qquad f = \frac{\gamma}{2\pi} B_0$$

where γ, the gyromagnetic ratio, is a known constant for each type of atom. For ^1H, $\gamma/2\pi = 42.58$ MHz/Tesla. As an analogy, precession of a nuclear spin about an external magnetic field is similar to a spinning top does in a gravitational field, see Fig. 1.5.

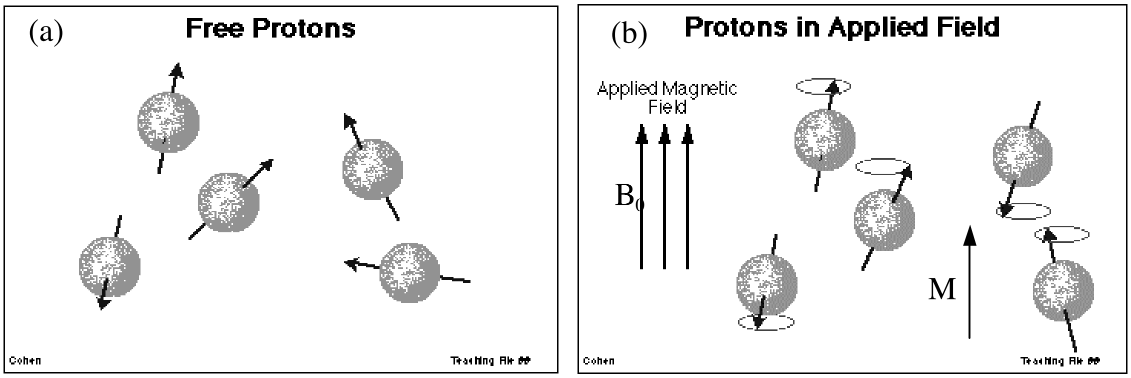

Fig. 1.4 Nuclear magnetic moment vectors (a) pointing in random directions and (b) aligned in the direction of an external magnetic field.

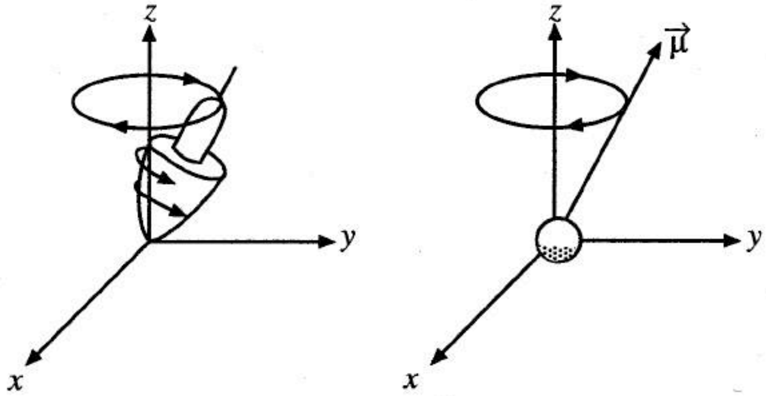

Fig. 1.5 Precession of a nuclear spin about an external magnetic field is similar to the wobbling of a spinning top in a gravitational field [2].

At equilibrium, the transverse component of M_0 (i.e., the projection of M_0 in the x-y plane, namely, M_{xy}) is zero (Fig 1.6a). To obtain MR signal, a RF field B_1 applies an external force the transverse plane to excite these spins out of equilibrium and tip M_0 away from the z-axis (the direction of B_0), creating a measurable (nonzero) transverse component M_{xy} (Fig. 1.6b). When turning the excitation off, relaxation back to its equilibrium also occurs with the length of the magnetization vector not remaining constant over time.

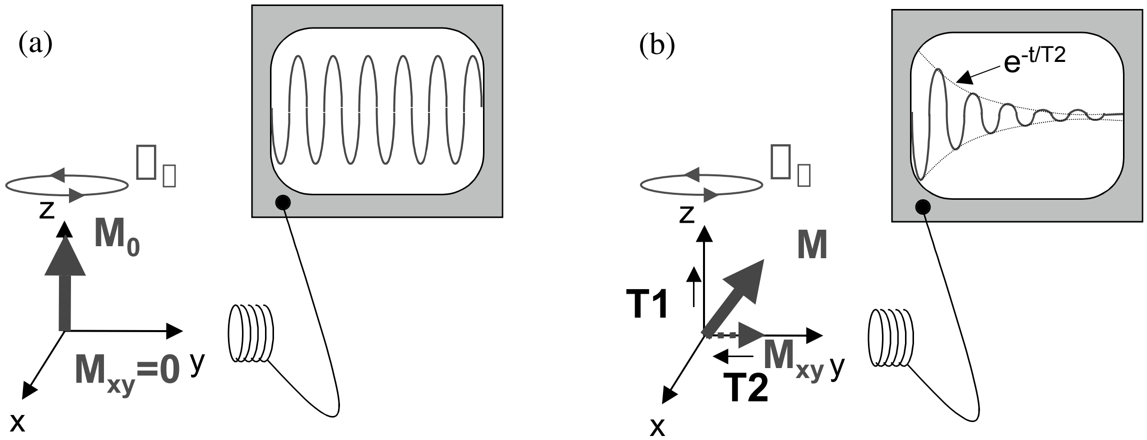

Fig 1.6 (a) The precessing nuclei induce a voltage in a receiver coil placed parallel to the x-y plane. Frequency of induced voltage is the same as the precessional frequency $\square_0$. (b) Detected signal is modulated by the fundamental relaxation processs, T_1 and T_2 decays. Both processes decrease M_{xy} as M $\square$ M_0.

The time constant characterizing the return of the magnetization vector along the z-axis (longitudinal axis) is called T_1, while the time constant characterizing the decay of the transverse component is called T_2 . In humans, T_1 values of most tissues range from about 100-1500 ms (Fig. 1.7), and its values increase as B_0 increases. T_2 values range from about 20-300ms (Table 1.1) and it is largely independent of field strength.

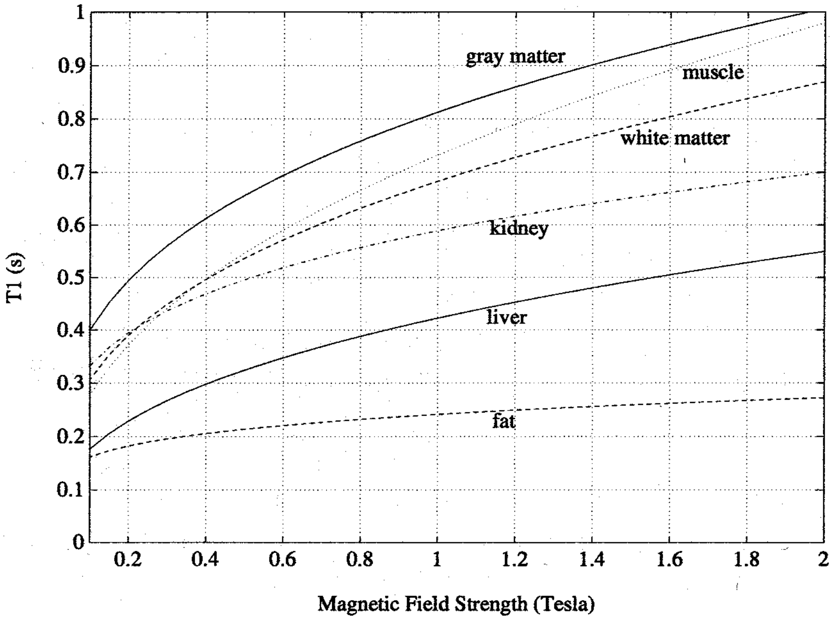

Fig. 1.7 Approximate T_1 values as a function of B_0 [1].

Tissue	T_2 (ms)
gray matter	100
white matter	92
muscle	47
fat	85
kidney	58
liver	43

Table 1.1 T_2 of some normal tissue types [1].

MR images are extremely rich in information content. The image pixel value can be considered as a function of a host of parameters, including the relaxation time constants T_1 and T_2, and the proton density (that has distinct values for different tissues). Therefore, by changing the effect of these parameters, MR images obtained from the same anatomical position can look drastically different (see Fig. 1.8). Note that left brain shown in an image is the right brain of the subject, and vise versa. These images can be further processed to produce new maps regarding water diffusion, blood flow, etc. Hence, the flexibility in data acquisition and the rich contrast mechanisms of MRI endow the technique with superior scientific and diagnostic values (Fig. 1.9).

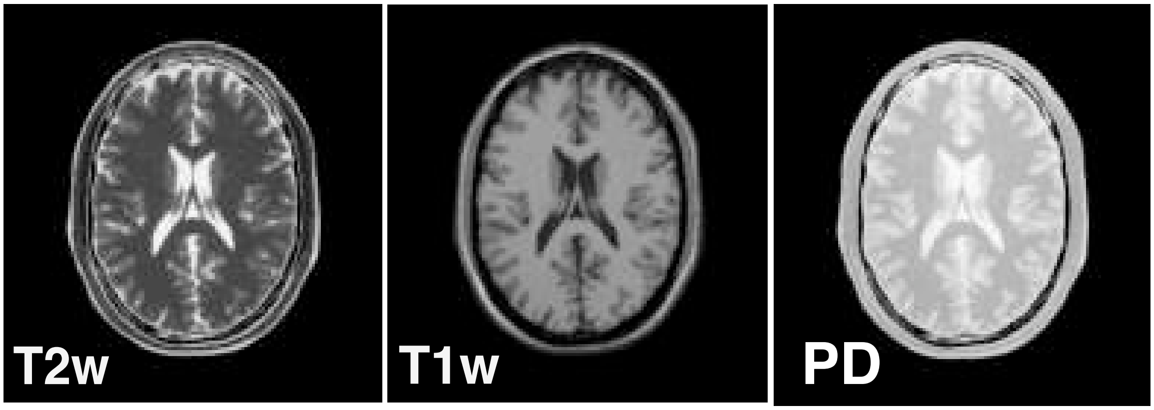

Fig. 1.8 Three images (a) T_2-weighted (T2w), (b) T_1-weighted (T1w), and (c) proton density (PD) images were obtained from the same cross section of a human head.

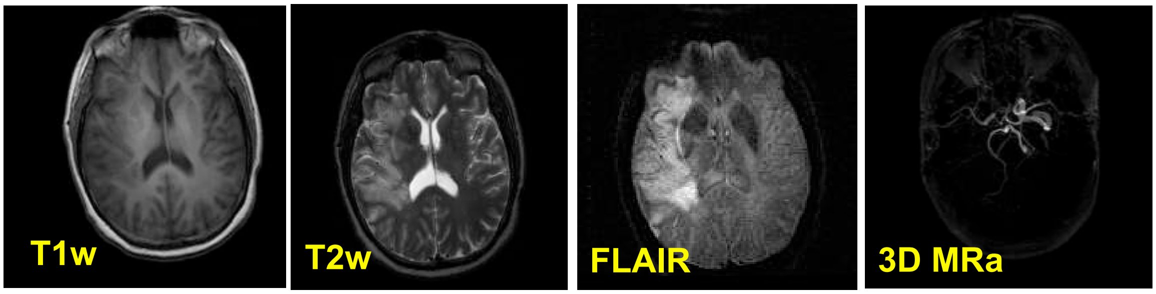

Fig. 1.9 Multiple MR images were acquired to help neurologists diagnose a stroke patient who had middle cerebral artery occlusion in comparison to the right side of the brain.

Further Reading:

1. DG Nishimura, Principles of Magnetic Resonance Imaging, April 1996.
2. Z-P Liang and PC Lauterbur, Principles of Magnetic Resonance Imaging: a signal processing perspective, IEEE press series in biomedical engineering, 1999.
3. F Bloch, “Nuclear induction,” Phys. Rev. vol 70, pp 460-474, 1946.
4. N Bloembergen, EM Purcell, and RV Pound, “Relaxation effects in nuclear magnetic resonance absorption,” Phys. Rev. vol. 73, pp 679-712, 1948.
5. Hornak’s on-line MR manual, www.cis.rit.edu/htbooks/mri/inside.htm
6. www.mri.swmed.edu/physwebold/mriphys1/index.htm
7. D Canet, Nuclear Magnetic Resonance: Concepts and Methods, John Wiley & Sons, New York, 1996
8. RH Hashemi and WG Bradley, Jr., MRI: The Basics, Williams & Williams, Baltimore, MD, 1997
9. M NessAiver, All You Really Need to Know about MRI Physics, Simply Physics, 5816 Narcissus Ave, Baltimore, MD, 1997
10. WJ Schempp, Magnetic Resonance Imaging: Mathematical Foundations and Applications, John Wiley and Sons, New York, 1998

2. FUNDAMENTALS OF THE FUNDAMENTALS: FOURIER THEORY

This section reviews Fourier theory, fundamentals for understanding MRI artifacts and image processing. Especially, we will focus on a few essential topics: the Fourier convolution theorem, sampling theory, and discrete Fourier transforms. A complete coverage can be found in [2,11,12].

Given a continuous function $f(x)$ of a real variable x, the Fourier transform of $f(x)$ is defined by

$$f(x) \xrightarrow{\text{Fourier transform}} F(u) = \int_{-\infty}^{+\infty} f(x) \exp(-j2\pi ux)\, dx$$

Conversely, give the Fourier response $F(u)$, $f(x)$ can be reconstructed by the inverse Fourier transform

$$f(x) = \int_{-\infty}^{+\infty} F(u) \exp(j2\pi ux)\, du \xleftarrow{\text{Inverser Fourier Transform}} F(u)$$

We usually call $f(x)$ and $F(u)$ a Fourier pair, both of which contain the same amount of information but just display the information in different spaces.

A two-dimensional Fourier transform is simply a sequence of two one-dimensional Fourier transforms, one along the x-axis and the other along the y-axis, namely

$$F(u,v) = \mathcal{F}_y\left\{\mathcal{F}_x\left\{f(x,y)\right\}\right\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\, exp\left(-j2\pi(ux+vy)\right) dxdy$$

Similarly, we can define the 2D inverse Fourier transform, suitably scaled, by

$$f(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(u,v)\, exp\left(j2\pi(ux+vy)\right) dudv$$

Some properties of the Fourier transform relevant to MRI are summarized below for easy reference.

a) Uniqueness: $f_1(x) = f_2(x) \Leftrightarrow F_1(u) = F_2(u)$

b) Linearity: $af_1(x) + bf_2(x) \Leftrightarrow aF_1(u) + bF_2(u)$

c) Shifting theorem:

$$f(x - x_0) \Leftrightarrow F(u)e^{-j2\pi ux_0}$$

$$e^{-j2\pi u_0 x} f(x) \Leftrightarrow F(u - u_0)$$

d) Conjugate symmetry: $f^*(x) \Leftrightarrow F^*(-u)$. Specifically, if $f(x)$ is a real-valued function, then $f(x) \Leftrightarrow F(u) = F^*(-u)$

e) Scaling property: $f(ax) \Leftrightarrow \frac{1}{|a|}F\left(\frac{u}{a}\right)$

f) Parseval's formula: the energy is conserved in both the space and frequency domains, i.e.,

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |F(u)|^2 du$$

g) Derivative property:

$$(-j2\pi x)^n f(x) \Leftrightarrow \frac{d^n F(u)}{du^n}$$

$$\frac{d^n f(x)}{dx^n} \Leftrightarrow (j2\pi u)^n F(u)$$

h) Convolution theorem:

$$f_1(x) \otimes f_2(x) \Leftrightarrow F_1(u)F_2(u)$$

$$f_1(x)f_2(x) \Leftrightarrow F_1(u) \otimes F_2(u)$$

Fourier Convolution Theorem:

It is worth spending more time to understand convolution because it is one of the core concepts in image processing. Mathematically, the convolution of two functions $f(x)$ and $g(x)$, denoted by $f(x) \otimes g(x)$, is defined as

$$f(x) \otimes g(x) = \int_{-\infty}^{+\infty} f(\tau) g(x - \tau) d\tau$$

Fig. 2.1 and 2.2 illustrate the concept of convolution graphically. From the image processing point of view, convolution can be regarded as a blurring or smoothing process; while deconvolution is a reverse process of convolution that undoes "blurring" and recovers the original information.

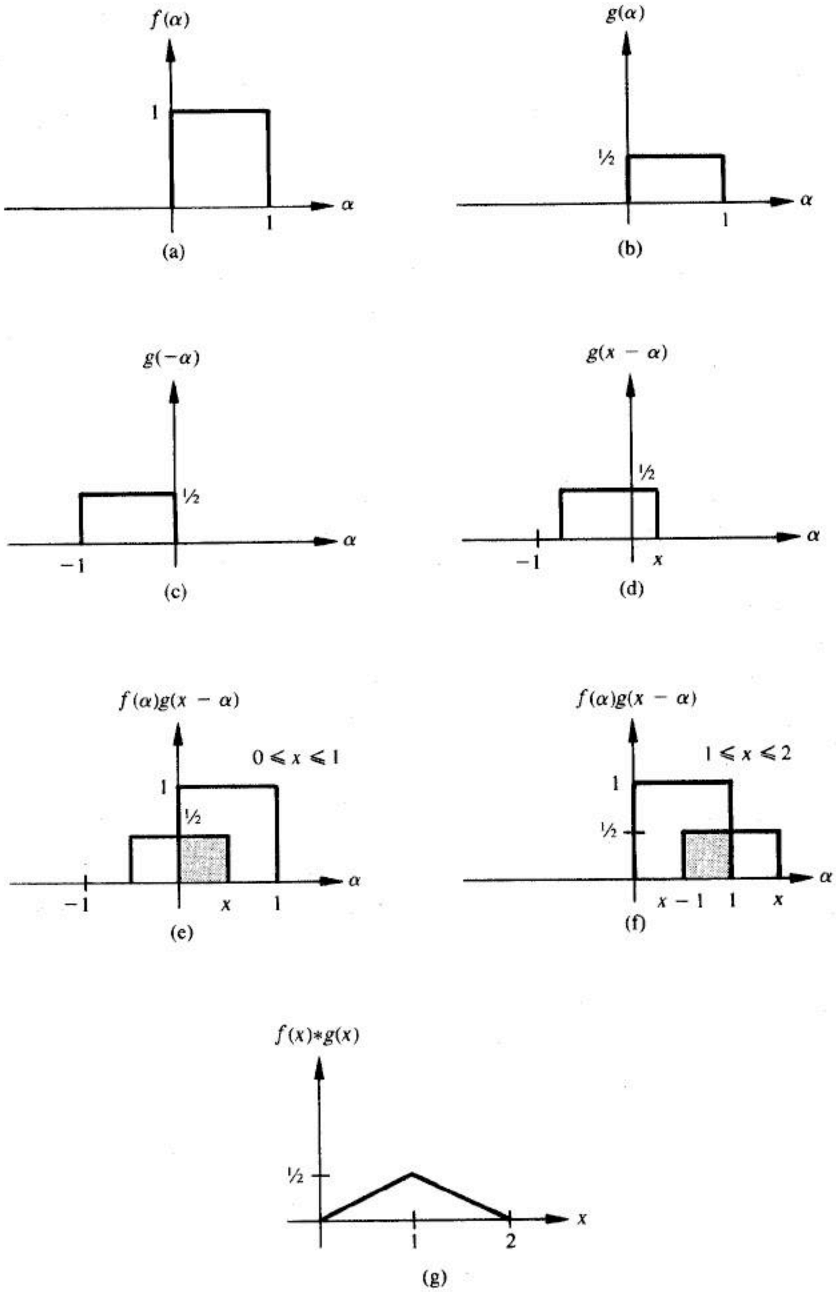

Fig. 2.1 Graphical illustration of convolution. The shaded areas indicate regions where the product is non-zero [11].

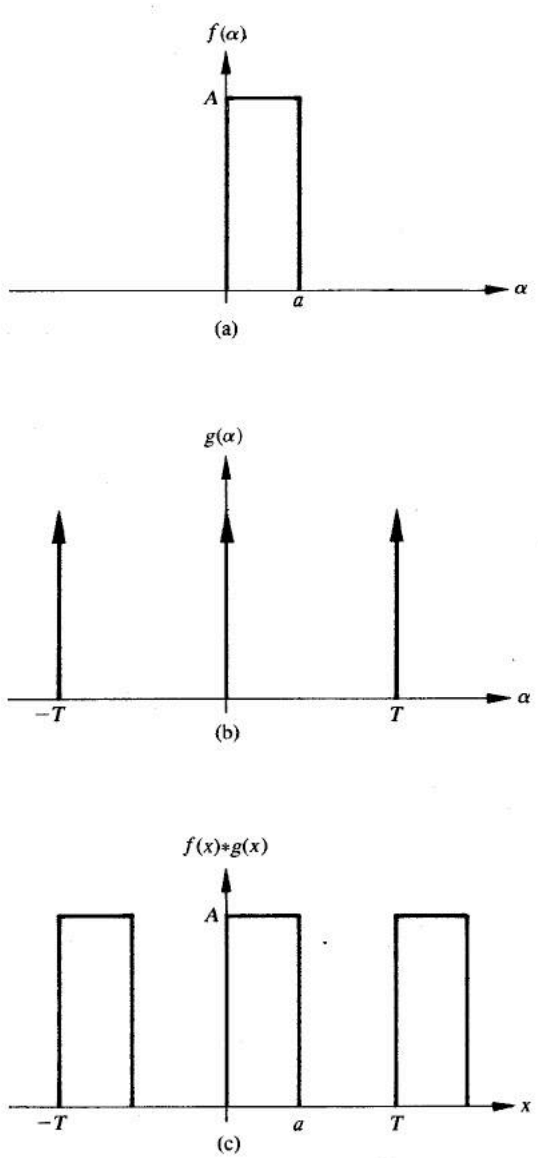

Fig. 2.2 Convolving with impulse functions [11].

As we can see, it is not easy to work out convolution by calculating the integral. Fortunately, the Fourier convolution theorem provides us a simple way to do so, i.e., we can accomplish convolution in one domain by doing a multiplication in another domain and inverse Fourier transforming back.

Sampling Theory:
Sampling in the time or object domain is a common scenario in many applications, however, in MRI, it is the opposite situation of sampling in the frequency domain. But the underlying sampling theory is the same. Here for simplicity, we will review the effects of finite sampling through Figs. 2.3 and 2.4 in 1D setting. Extension to the 2D case is straightforward.

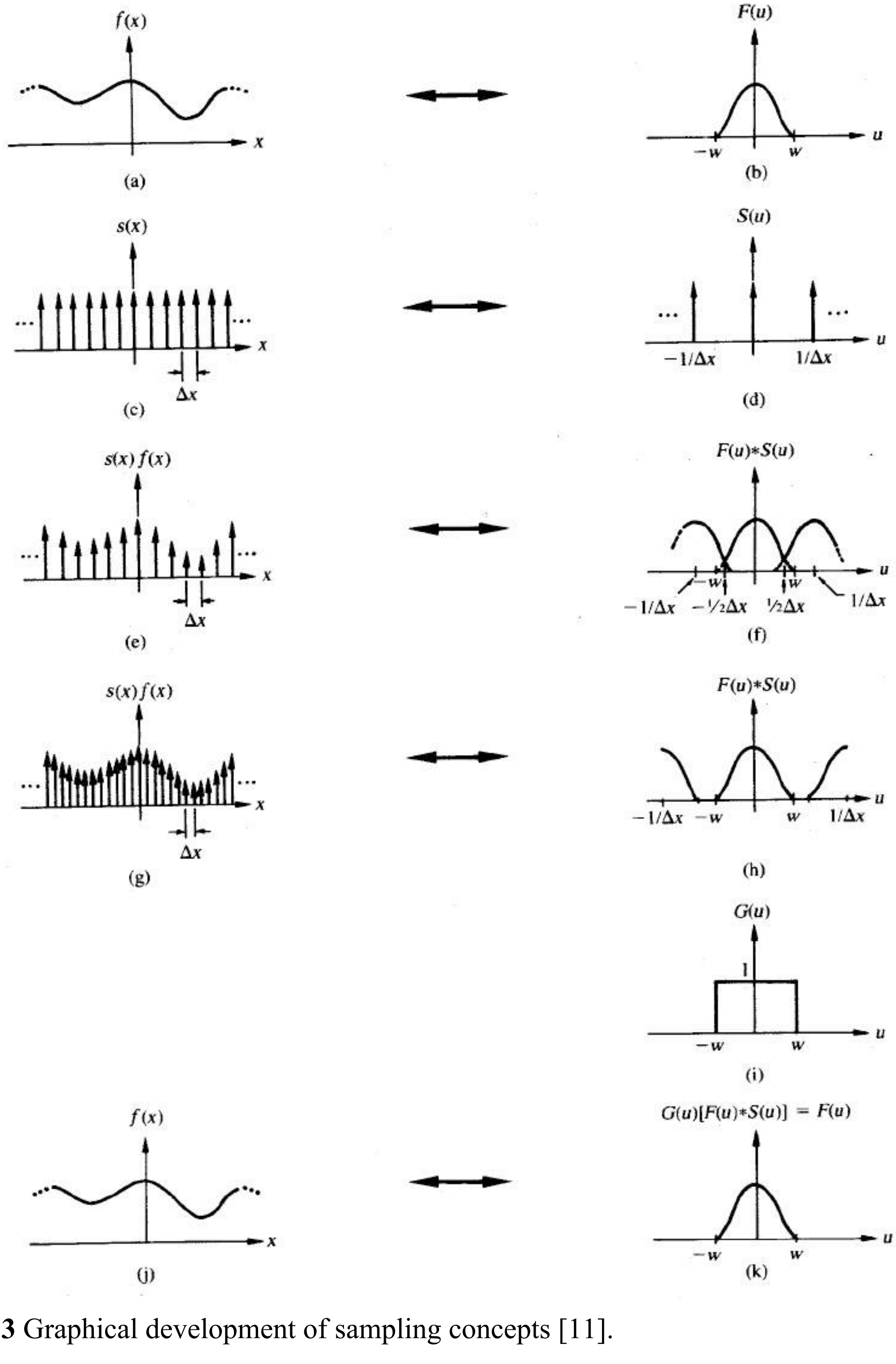

Fig. 2.3 Graphical development of sampling concepts [11].

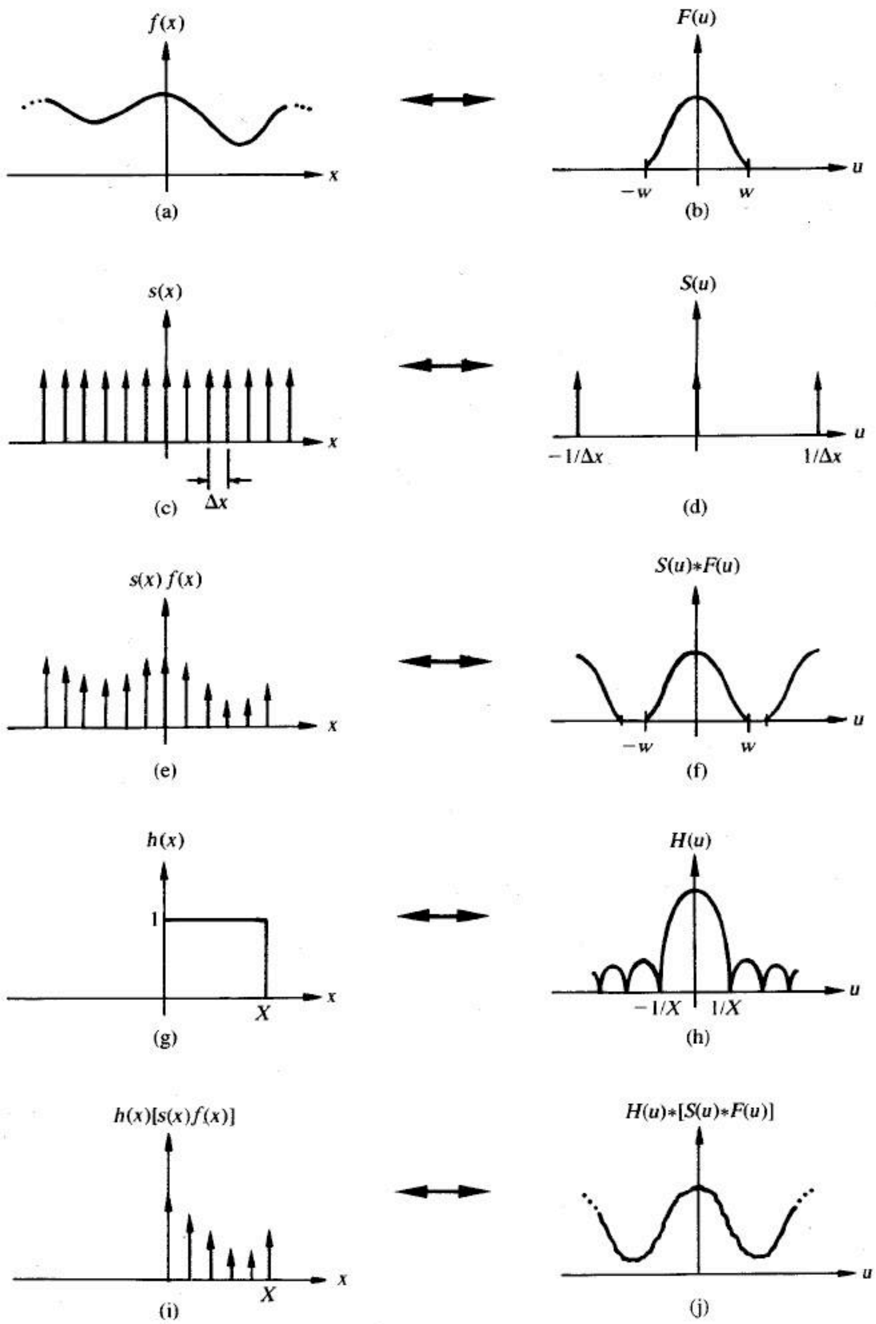

Fig. 2.4 Graphical illustration of finite-sampling concepts [11].

More importantly, one may wonder if all information can be reserved when sampling data. In fact, this can be true under certain condition. A function $f(x)$ is “band-limited” if $F(u)$ is of finite extent, i.e., all the frequency information of a band-limited function is contained in the finite interval [-W, W]. A band-limited function can be recovered completely from samples whose spacing, Δx, satisfies

$$\Delta x \leq \frac{1}{2W}$$

This is the *Shannon sampling theorem*. Note that if the sampling period is in units of cm^{-1}, then the sampling rate in the frequency domain is in units of cm. If the sampling rate is too low, then the replications overlap, resulting in a condition known as "*aliasing*" (see Fig 2.3f). The minimum sampling rate that avoids aliasing is called the "*Nyquist rate*" and is equal to twice the maximum frequency of $f(x)$, i.e., 2W.

Discrete Fourier Transform:

Sampling theory provides a means to understand the periodicity of the discrete Fourier transform, see Fig. 2.5. One keeps periodicity in mind when processing signals through the discrete Fourier transform.

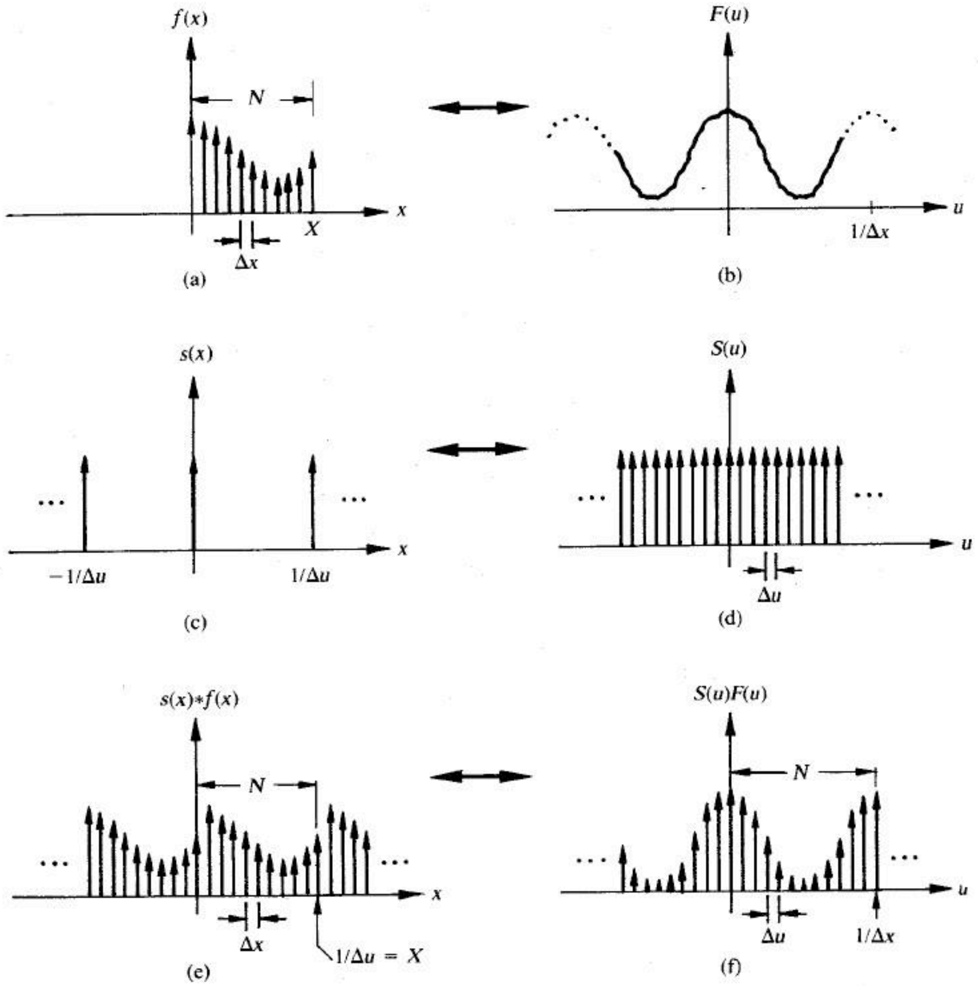

Fig. 2.5 Graphical illustration of the discrete Fourier transform [11].

Further Reading:

11. RC Gonzalez and P Wintz, "Digital Image Processing", second edition, Addison-Wesley, 1987

12. RC Gonzalez and RE Woods, "Digital Image Processing", third edition, Addison-Wesley, 1993
13. RJ Marks II, Introduction to Shannon Sampling and Interpolation Theory, Springer-Verlag, New York, 1991
14. RJ Marks II, Advanced Topics in Shannon Sampling and Interpolation Theory, Springer-Verlag, New York, 1993

3. COMMON MR IMAGE ARTIFACTS

The Fourier transform and the sampling theory can help us to understand some common artifacts appeared in MR imaging process.

Gibbs Ringing Artifacts:

The Gibbs ringing artifact, a common image distortion that exists in Fourier images, manifests itself as spurious ringing around sharp edges, as illustrated in Fig 3.1. This is a result of truncating the Fourier spectrum due to finite sampling or missing of high-frequency data.

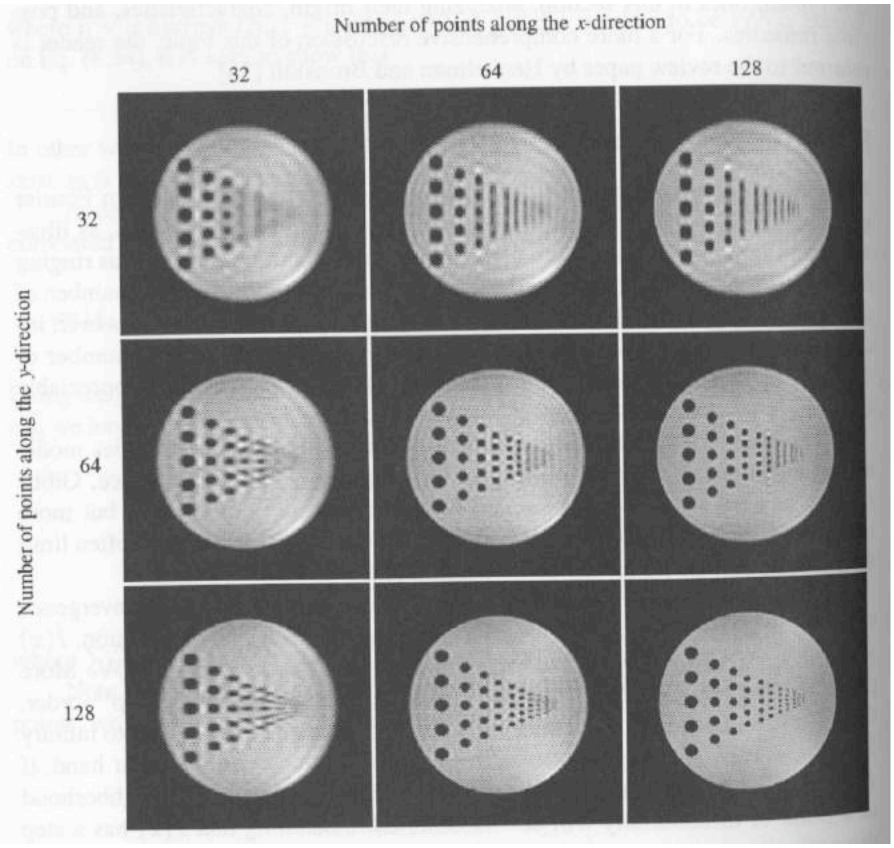

Fig 3.1 Gibbs ringing artifacts in Fourier reconstructions. Notice the ringing pattern as a function of the number of data points used in the reconstruction [2].

Aliasing Artifacts:

When the sampling rate is less than the Nyquist rate, perfect reconstruction is not possible and the resulting error creates aliasing artifacts, where replications of the object overlap (Fig. 3.2). The appearance of aliasing artifacts is dependent on how sampling is done. Aliasing artifacts become less structured and thus less obvious for non-uniform sampling. This property is sometimes utilized to reduce the aliasing artifacts when signal undersampling cannot be avoided.

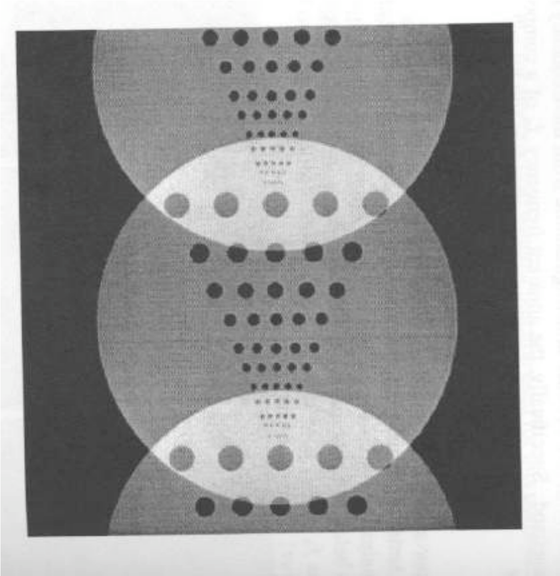

Fig 3.2 Aliasing artifacts due to undersampling along the vertical direction [2]

<u>Motion artifacts</u>:
In practice, the object being imaged is not usually stationary during the data acquisition period. As a consequence, image artifacts result. The most problematic physiologic motions include blood flow, respiratory motion, cardiac motion, and gross movements of the body. Common motion artifacts are image blurring and ghost (or misregistration) as shown in Fig 3.3.

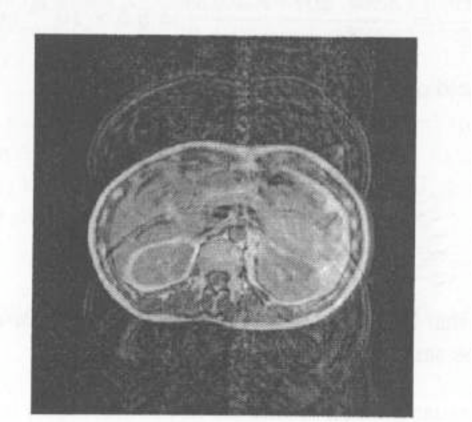

Fig 3.3 Cross-sectional image of the lower abdomen with motion artifacts [2].

Further reading

15. RM Henkelman and MJ Bronskill, “Artifacts in magnetic resonance imaging”, Rev. Magn. Reson. Med., vol 2, pp1-126, 1987

16. S Amartur and EM Haacke, “Modified iterative model based on data extrapolation method to reduce Gibbs ringing,” J. Magn. Res. Imag., vol 1 pp307-317, 1991
17. D Atkinson, DLG Hill, PN Stoyle, PE Summers, and SF Keevil, “Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion,” IEEE Trans. Med. Imag. vol 16, pp903-910, 1997
18. JL Duerk and OP Simonetti, “Review of MRI gradient waveform design methods with application in the study of motion”, Concepts Magn. Reson., vol 5, No 2, pp105-122, 1993
19. J Jackson, C Meyer, D Nishimura, and A Macovski, “Selection of a convolution function for Fourier inversion using gridding”, IEEE Trans. Med. Imag. Vol MI-10, pp473-478, 1991
20. V Rasche, R Proksa, R Sinkus, P Bornert, and H Eggers, “Resampling of data between arbitrary grids using convolution interpolation,” IEEE Trans. Med. Imag., vol 18, pp. 385-392, 1999
21. QS Xiang and RM Henkelman, “K-space description for MR imaging of dynamic objects,” Magn. Reson. Med., vol 29, pp. 422-428, 1993
22. QS Xiang, MJ Bronskill, and RM Henkelman, “Two-point interference method for suppression of ghost artifact due to motion,” J. Magn. Reson. Imag. Vol. 5, pp. 529-534, 1995
23. X Zhou, ZP Liang, GP Cofer, CF Beaulieu, SA Suddarth, and GA Johnson, “Reduction of ringing and blurring artifacts in fast spin-echo imaging,” J. Magn. Reson. Imag., vol 3, pp. 803-807, 1993

4. COMPUTER-AIDED DIAGNOSTIC PROCESSING

Medical image processing deals with the development of problem-specific approaches to the enhancement of raw medical image data for the purposes of selective visualization as well as further analysis. There are many topics in medical image processing: some emphasize general applicable theory and some focus on specific applications. A comprehensive overview on a broader range of topics in medical image processing appears in [24]. Here, we mostly focus on image segmentation and multi-spectral analysis.

<u>Image segmentation</u>:

Image segmentation is defined as a partitioning of an image into regions that are meaningful for a specific task; it is a labeling problem. This may, for instance, involve the detection of a brain tumor from MR or CT images (Fig. 4.1). Segmentation is one of the first steps leading to image analysis and interpretation. The goal is easy to state, but difficult to achieve accurately.

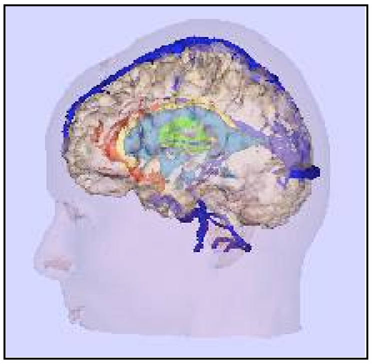

Fig. 4.1 A 3D rendering of segmented skin surface (pink), brain tissue (brown), major blood vessels (navy blue), and a tumor (green) from MRI volume. This allows surgeons to visualize the actual location and to plan and simulate specific procedures.

Image segmentation approaches can be classified according to both the features and the type of techniques used. Features include pixel intensities, edge information, and texture, etc. Techniques based on these features can be broadly classified into structural and statistical methods.

Structural methods are based on the spatial properties of the image, such as edges and regions. Various edge detection algorithms have been applied to extract boundaries between different brain tissues [26-28]. However such algorithms are sensitive to artifacts and noise. Region growing [29,30] is another popular structural technique. In this approach, one begins by dividing an image into small regions, which can be considered as ``seeds''. Then, all boundaries between adjacent regions are examined. Strong boundaries (in terms of certain specific properties) are kept, while weak boundaries are rejected and the adjacent regions merged. The process is carried out iteratively until no boundaries are weak enough to be rejected. However, the performance of the method depends on seed selection and whether the regions are well defined, and therefore is also not considered robust.

Starting from a totally different viewpoint, statistical methods label pixels according to probability values, which are determined based on the intensity distribution of the image.

Gray-level thresholding is the simplest, yet often effective, segmentation method. In this approach, structures in the image are assigned a label by comparing their gray-level value to one or more intensity thresholds. A single threshold serves to segment the image into only two regions, a background and a foreground, as illustrated in Fig. 4.2. Sometimes the task of selecting a threshold is quite easy, when there is a clear difference between the gray-levels of the objects we wish to segment.

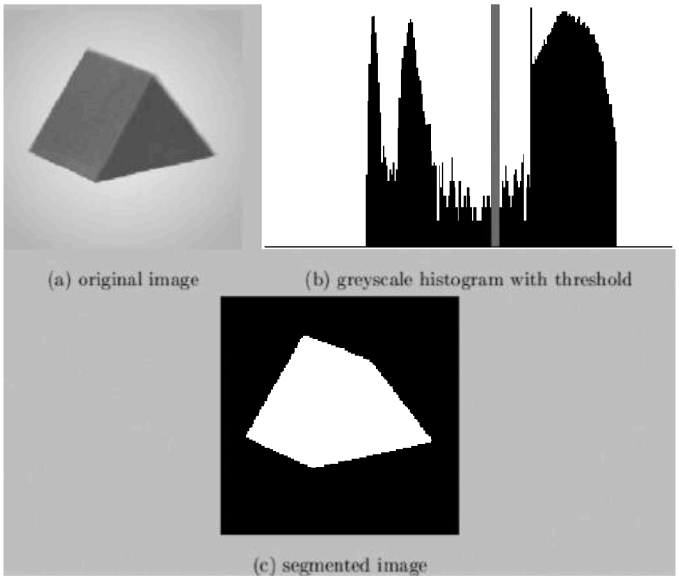

Fig 4.2 Segmenting a simple image by a single threshold [25].

However, things are not normally so simple. Inhomogeneity in the imaging equipment and the partial-volume effect (multiple tissue class occupation within a voxel) give rise to a smoothly varying, non-linear gain field. While the human visual system easily compensates for this field, the gain can perturb the histogram distributions, causing significant overlaps between intensity peaks and thus leading to substantial misclassification in traditional intensity based classification methods.

Hence, there are more sophisticated statistical approaches [31-39] relying on certain assumptions or models of the probability distribution function of the image intensities and its associated class labels, which can both be considered random variables. Let C and Y be two random variables for the class label and the pixel intensity, respectively, and c and y be typical instances. The class-conditional density function is p(y|c). Statistical approaches attempt to solve the problem of estimating the associated class label c, given only the intensity y for each pixel. Such an estimation problem is necessarily formulated from an established criterion. Maximum a posteriori (MAP) or maximum likelihood (ML) principles are two such examples. Many statistical segmentation methods differ in terms of models of p(y). Depending on whether a specific functional form for the density model is assumed, a statistical approach can either be *parametric* or *non-parametric*. Both have been widely used in segmentation of brain MR images.

In non-parametric methods, the density model p(y) relies entirely on the data itself, i.e. no prior assumption is made about the functional form of the distribution but a large number of correctly labeled training points are required in advance.

One of the most widely used non-parametric methods is the K-Nearest-Neighbors (KNN) rule classification. First fix K, the number of nearest neighbors to find in the neighborhood of any unlabelled intensity y. The KNN involves finding a neighborhood around the point y that contains K points independent of their class, and then assigning y to the class having the largest number of representatives inside the neighborhood [38, 39]. When K = 1, it becomes the nearest-neighbor rule, which simply assigns a point y to the same class as that of the nearest point from the training set.

Non-parametric methods are adaptive, but suffer from the difficulty of obtaining a large number of training points, which can be tedious and a heavy burden even for experienced people. Clearly, such methods are not fully automatic.

Unlike non-parametric approaches, parametric approaches rely on an explicit functional form of the intensity density function. For instance, the intensity density function can be modeled by a sum of Gaussian distributions, each of which models an intensity distribution of each class. Here, means and variances of Gaussian distributions becomes parameters of the model to be determined.

Maximum likelibood (ML) methods aim at estimating a set of parameters that maximize the probability of observing the gray-level distribution. A good method to estimate these parameters is to use an Expectation-Maximization (EM) approach, an iterative technique. Figures 4.3 and 4.4 show examples of results obtained with the EM algorithm on real MR data.

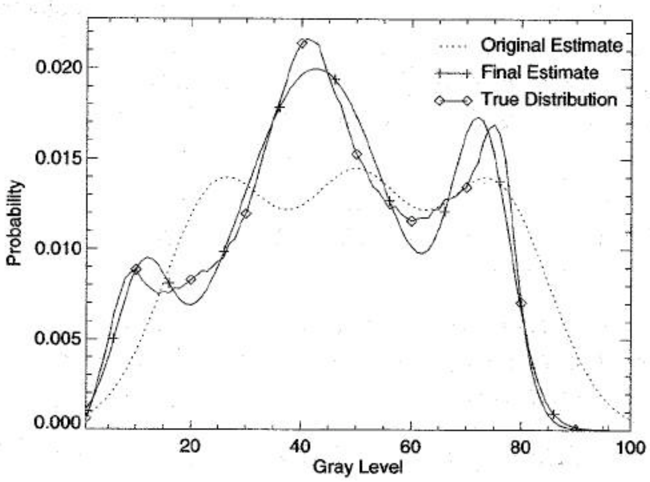

Fig 4.3 EM algorithm for parameter estimation. The histogram was computed from real MR data of a brain. The three peaks correspond to cerebral spinal fluid (CSF), gray matter (GM), and white matter (WM), respectively [24].

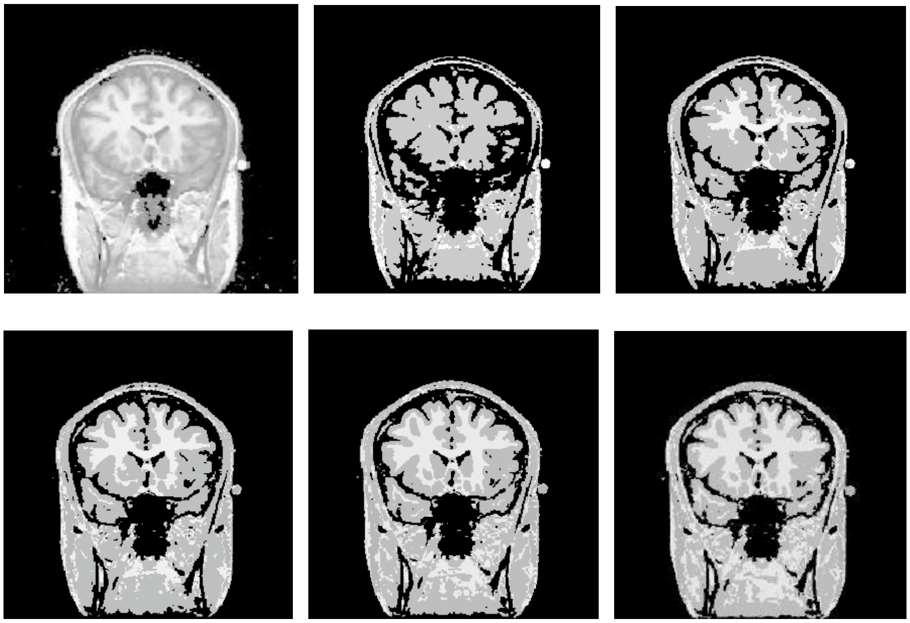

Fig. 4.4 Top to bottom, left to right: Input image and tissue classification generated by successive iterations of the EM segmentation (WM is brightest, GM is medium grey, and CSF and air are black). Bottom right is the final segmentation.

Multi-spectral Analysis:

We have talked about image processing based on one single MR image. However, MRI data are by nature multi-spectral because their contrast characteristics depend on the acquisition sequences and their parameters. As an example, Fig 4.5 shows proton density weighted, T2-weighted, T1-weighted, and gadolinium-DTPA enhanced T1w (GAD) brain MRI scans of the same multiple sclerosis (MS) patient. Extracting data from multi-spectral MR exams may provide us additional useful information in disease diagnosis and treatment assessment [40].

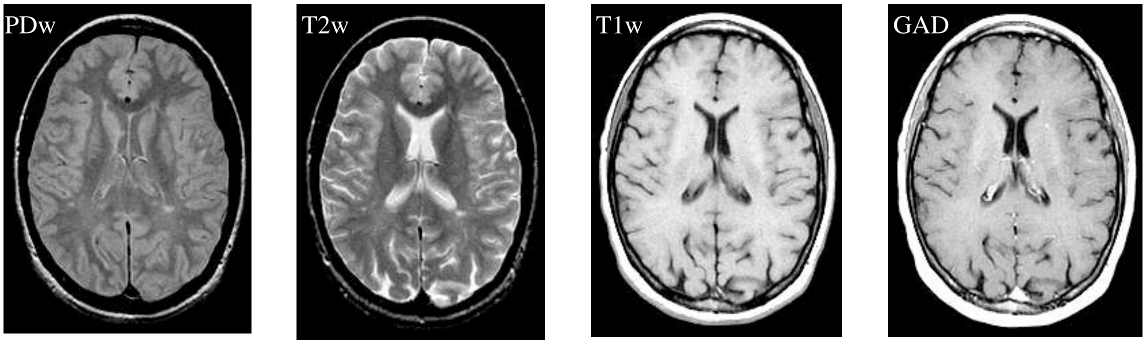

Fig. 4.5 Example of multi-spectral MR images. From left to right: proton density weighted, T2-weighted, T1-weighted, and gadolinium-DTPA enhanced T1w (GAD) brain MRI scans of the same MS patient [40].

We aim to improve differentiation of lesions (abnormal tissues) and normal brain tissues using the information provided by multi-spectral MRI. [40] describes a medical application of a method used successfully in the past to extract and summarize information from satellite remote sensing data [41]. The technique acquires multiple co-registered images, and constructs a “spectrum” for each pixel from its intensity values in each image, see Fig. 4.6. Then we can produce synthetic images of "spectral phase" (SP) relative to a reference tissue, by calculating the angular difference between any pixel spectrum and that of the reference tissue, as illustrated in Fig. 4.7. These SP images allow retrospective suppression of signals in the reference tissue, while improving tissue contrast between different types of tissues, as shown in Fig 4.8. The additional contrast in SP images not only may aid the quantification and analysis of lesion activity in MR exams of MS patients, but also can be further used to improve segmentation performance.

Multi-spectral analysis has a great potential in medical image analysis, tissue segmentation and classifications [42-45].

There are many other exciting research projects going on worldwide. You may do a google search to find out more. Also, you may refer to [46] for on-going research in the Seaman Family MR Research Centre in Calgary, Alberta.

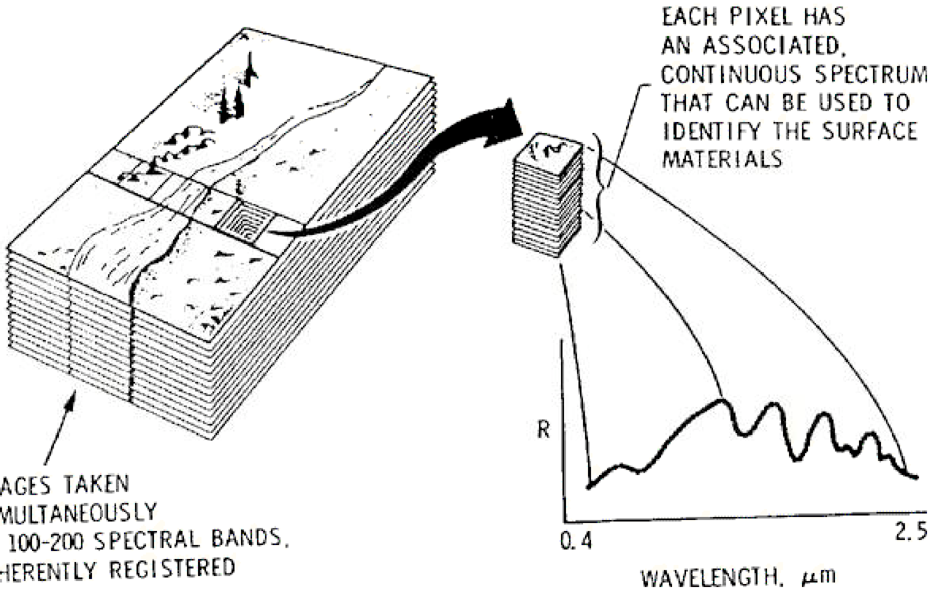

Fig 4.6 Each image pixel is associated with a spectral vector constructed from its intensities in each image, where all images are co-registered.

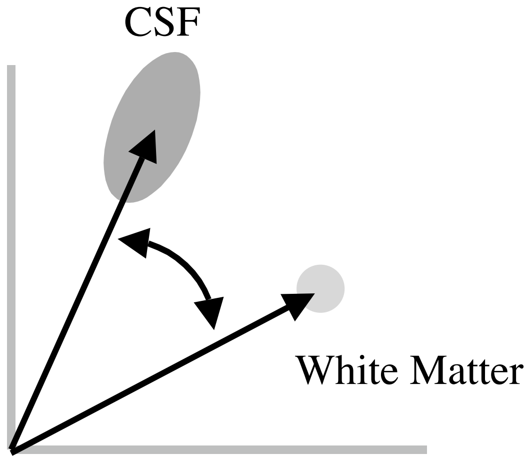

Fig 4.7 For each pixel, calculate the angle from this vector to the reference vector.

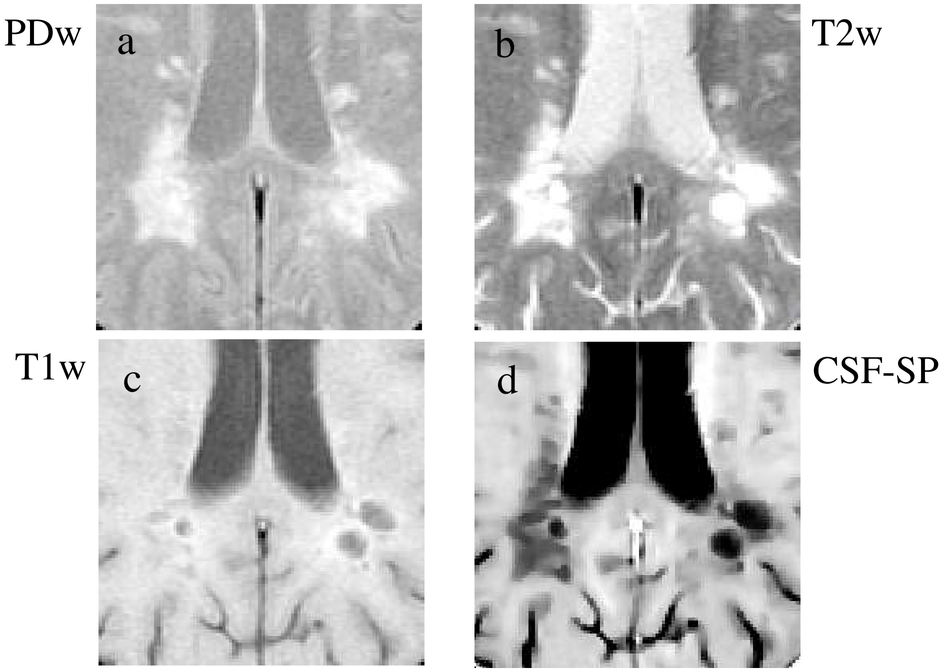

Fig 4.8 Magnified views of a large MS lesion in an MS patient. a) through c) are the peri-ventricular lesion in the PDw, T2w and T1w images, respectively. d) shows the same lesion in the derived spectral-phase image relative to CSF (CSF-SP) [40].

Further Reading:

24. M Sonka, JM Fitzpatrick, Handbook of Medical Imaging: volume 2 Medical Image Processing and Analysis, SPIE, Washington, 2000
25. Segmentation, http://www.redbrick.dcu.ie/~bolsh/thesis/node25.htm
26. M Bomans, KH Hohne, LA Kramer, and JM Fletcher, "3-D segmentation of MR images of the head for 3-D display," IEEE Trans. Med. Imag., 18(1):25-34, 1990
27. LP Clarke, RP Velthuizen, MA Camacho, JJ Heine, M Vaidyanathan, LO Hall, RW Thatcher, ML Silbiger, "MRI segmentation: methods and applications", Mag. Reson. Imag., vol 13, no. 3 pp. 343-368, 1995
28. S Dellepiane, "Image segmentation: Errors, sensitivity, and uncertainty," Proc. 13th IEEE-Eng. Med. Bio. Society, vol. 13, pp. 253-254, 1991
29. CR Brice and CL Fennema, "Scene analysis using regions," Artificial Intelligence, vol. 1, No. 3, pp. 205-226, 1970
30. SC Zhu and A Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation," IEEE Trans. Pattern Anal. Machine Intell., vol. 18, No. 9, pp. 884-900, 1996
31. P Taylor, "Invited review: computer aids for decision-making in diagnostic radiology—a literature review," Brit. J. Radiol., vol. 68, pp. 945—957, 1995
32. CA Glasbey, "An analysis of histogram-based thresholding algorithms," CVGIP—Graphics Models and Image Processing, vol. 55, pp. 532—537, 1993
33. P Sahoo, S Soltani, A Wong, Y Chen, "Survey of thresholding techniques," Computer Vision, Graphics and Image Processing, vol 41, No.2 pp.233-260, 1988

34. A Dempster, N Laird, D Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Royal Statistical Society, series B, vol. 39, no. 1, pp.1-38, 1977
35. WM Wells III, WEL Grimson, R Kikinis, FA Jolesz, "Adaptive segmentation of MRI data," IEEE Trans. On Medical Imaging, vol. 15, pp.429-442, Aug. 1996
36. R Guillemaud, M Brady, "Estimating the bias field of MR images," IEEE Trans. On Med. Imag., vol 16, pp. 238-251, 1997
37. KV Leemput, F Maes, D Vandermeulen, P Suetens, "Automated model-based bias field correction of MR images of the brain," IEEE Trans. Med. Imag. 1999
38. JC Bezdek, LO Hall, LP Clarke, "Review of MR image segmentation techniques using pattern recognition," Medical Physics, vol. 20, No. 4 pp.1033-1048, 1993
39. JR Mitchell, SJ Karlik, DH Lee, A Fenster, "Computer-assisted identification and quantification of multiple sclerosis lesions in MR imaging volumes in the brain," J. Magn. Reson. Imag., vol. 4 pp.197-208, 1994
40. JR Mitchell, C Jones, SJ Karlik, K Kennedy, DH Lee, BK Rutt, A Fenster, " MR Multispectral Analysis of Multiple Sclerosis Lesions," J Magn. Reson. Imag., vol 7 No. 3, pp. 499-511, 1997
41. The environment for visualization images (ENVI), version 3.0 Users Guide, Better Solutions Consulting, Lafayette Colorado, 1997 edition
42. B Johnston, MS Atkins, "Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI,", IEEE Trans. Med. Imag., vol. 15, no. 2, pp. 1-15, 1996
43. RAD Carano, K Takano, KG Helmer, T Tatlisumak, K Irie, JD Petrucceli, M Fisher, CH Sotak, "Determination of focal ischemic lesion volume in the rat brain using multispectral analysis,", JMRI, vol. 8, pp. 1266-1278, 1998
44. H S-Zadeh, JP Windham, DJ Peck, "Optimal linear transformation for MRI feature extraction," IEEE Trans. Med. Imag., Vol. 15, no. 6, pp. 749-766, 1996
45. M Kamber, R Shinghal, DL Collins, GS Francis, AC Evans, "Model-based 3D segmentation of multiple sclerosis lesions in magnetic resonance brain images," IEEE Trans. Med. Imag., vol. 14, no. 3, pp. 442-453, 1995
46. Mitchell JR, Frayne R, Sutherland G. "Medical physics research in the acute care setting at the Seaman Family MR Research Center," Phys in Canada, vol. 58, No. 2, pp. 127-134, 2002

5. VALIDATION OF ALGORITHMS

Receiver operating characteristic (ROC) analysis is a common means of comparing precision, accuracy, and efficiency of various algorithms. ROC analysis can be effective for assessing how good a test is at discrimination between patients with the disease from those without the disease. Here, we can extend it to evaluate how good an algorithm (for instance, segmentation and classification algorithms) is for differentiating abnormal tissue from normal tissues. To keep it simple, we will use a diseased population as an example to explain ROC analysis.

Fig. 5.1 graphically illustrates the prevalence of disease and introduces the concepts of true negative (TN), false negative (FN), false positive (FP) and true positive (TP).

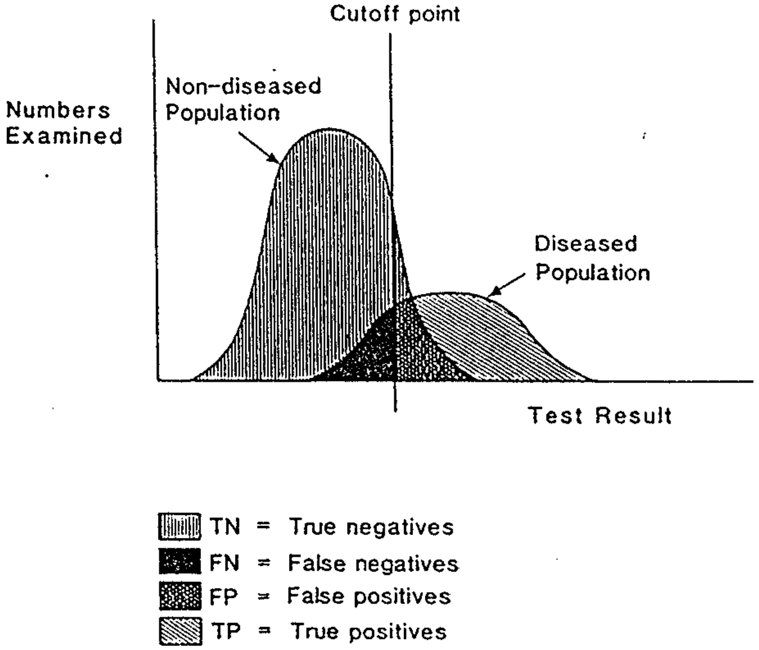

Fig. 5.1 This shows the overlap with FP and FN results when the prevalence in the population studied is less than 50%, i.e., there are a fewer diseased than normal people [47].

For any cutoff value of the diagnostic criterion used to separate diseased from non-diseased populations, we can express the outcome as a two by two matrix:

		TRUTH	
		P	**N**
TEST	**P**	TP	FP
	N	FN	TN

The accuracy of a test is defined as

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100\%$$

It is a global measure of accuracy and doesn't distinguish between positives and negatives. It may be misleading and is rarely of any value.

The sensitivity of a test is the measurement referred to as "the proportion of people with the disease who will be detected by the test", namely,

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

The specificity of a test is the measurement referred to as “the proportion of people without the disease who will be confirmed to be free of the disease by the test”, namely,

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%$$

When the cutoff point, i.e. threshold, is changed, sensitivity and specificity of the test vary too. High sensitivity picks up most of the abnormals and is good for excluding disease, screening, associated with missing the disease. High specificity picks up of the normals and is good for confirming the presence of disease and cases when there is high risk associated with treating unnecessary. A perfect test is both highly sensitive and highly specific.

A useful measure of the performance of a diagnostic test is the receiver operating characteristic (ROC) curve because it measures the sensitivities and specificities over a wide range of thresholds. These curves will demonstrate graphically how the sensitivity and the specificity change when the diagnostic criterion is relaxed or strict. Conventionally, levels of likelihood of disease being present are divided into five categories: 0, definitely normal; 1, probably normal; 2, possibly abnormal; 3, probably abnormal; 4, definitely abnormal.

The evaluation of a test will use these different levels of likelihood included in a report. From these data the ROC curves can be constructed for each test to express the performance of the test quantitatively. It is plotted in the ROC space, i.e. sensitivity *vs.* (1 - specificity), or, TP ratio *vs.* FP ratio, as shown in Fig. 5.2. These curves can be used to evaluate the overall performance of different tests (Fig. 5.3).

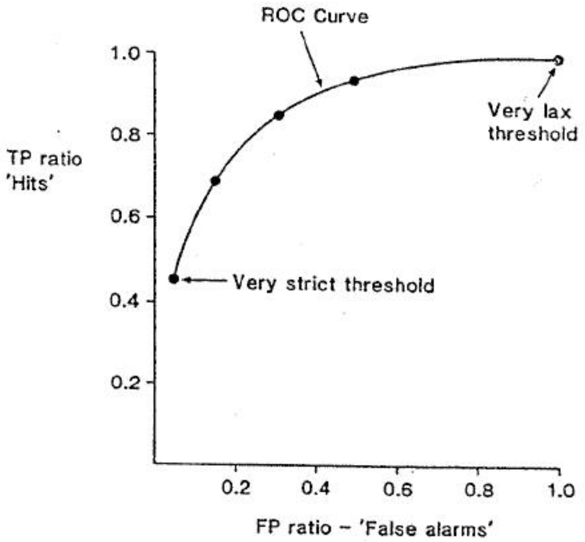

Fig. 5.2 An ROC curve from an imaging procedure with five level of diagnostic probability [47].

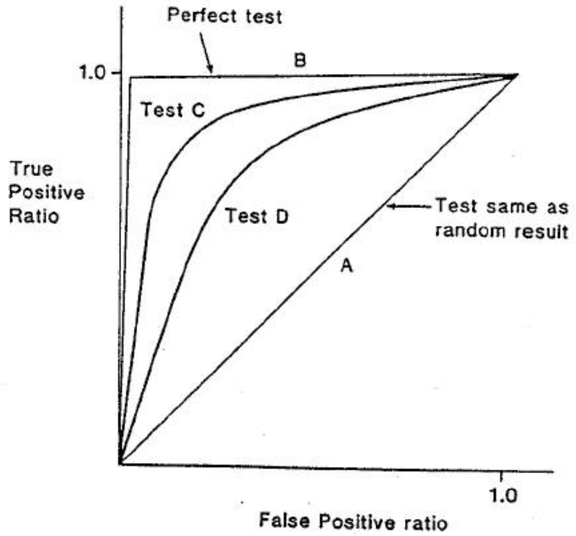

Fig. 5.3 Four theoretical ROC curves showing test performance from perfect (B) to random (A). C and D are typical for many imaging procedure [47].

The performance of the tests can be assessed by a single value, commonly the area below the curve denoted as A_Z. The closer to 1.0 the A_Z, the better the test performance. However, the A_Z measure may not be enough to distinguish ROC curves that cross over as shown in Fig. 5.4.

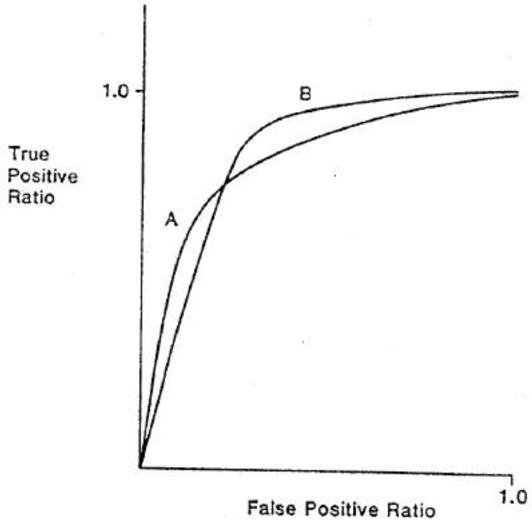

Fig. 5.4 An example of an ROC curve when the areas are similar but performance differs in different areas of the curve [47].

Conventional ROC analysis is strictly applicable to binary decision, abnormal or normal. Often clinical tasks does not fit binary model. For example, lesion segmentation reports contain location information (lesions may occur in multiple locations) as well as the binary diagnosis. To overcome this limitation, there are variants of ROC analysis in which location information is considered. In free-response ROC (FROC) analysis [53-56] the reader rates all locations on an image perceived as being suspicious. That is, for each image, the reader could yield zero or more ratings corresponding to these locations, and each image could contain zero (i.e., normal) or more lesions (i.e. abnormal). Localizations within a clinically acceptable distance from a true lesion are scored as TP events, and all other events are scored as FP events.

Further Reading

47. MN Maisey, J Hutton, “Guidelines for the evaluation of radiological technologies”, report of a working party of the British Institute of Radiology
48. CE Metz, “Basic principles of ROC analysis,” Seminars In Nuclear Medicine, vol. 8, pp.283-98, 1978
49. CE Metz, “ROC methodology in radiologic imaging,” Investigative Radiology, vol. 21, pp. 720-733, 1986
50. JA Hanley, “Receiver operating characteristic (ROC) methodology: The state of the art,” Critical Reviews in Diagnostic Imaging, vol. 29, pp. 307-335, 1989
51. YT van der Schouw, AL Verbeek, SH Ruijs, “Guidelines for the assessment of new diagnostic tests,” Invest Radiol, vol. 30, pp. 334-340,1995
52. JB Tilbury, PWJ Van Eetvelt, JM Garibaldi, JSH Curnow, EC Ifeachor, “Receiver operating characteristic analysis for intelligent medical systems—a new approach for finding confidence intervals,” IEEE Trans. Biomed. Eng., vol. 47, no. 7, pp. 952-963, 2000
53. DP Chakraborty, “Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data,” Medical Physics, vol. 16, pp. 561-568, 1989
54. D.P. Chakraborty, L.H. Winter, “Free-response methodology: Alternate analysis and a new observer-performance experiment,” Radiology, vol. 174, pp. 873-81, 1990
55. DP Chakraborty, “Statistical power in observer-performance studies: comparison of the receiver operating characteristic and free-response methods in tasks involving localization,” Aca. Radiol., vol. 9, pp. 147-156, 2002
56. MA Anastasio, MA Kupinski, RM Nishikawa, “Optimization and FROC analysis of rule-based detection schemes using a multiobjective approach,” IEEE Trans. Med. Imag., vol. 17, no. 6, pp. 1089-1093, 1998

6. SUMMARY

We hope that this session may help you learn the fundamentals of medical image processing, experience it as a plastic blend of science and art, and most importantly understand the ultimate goal in medical image processing--helping patients.