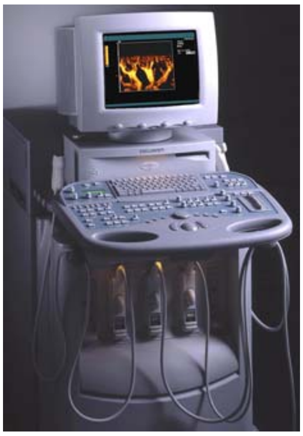
超音波生物醫學影像

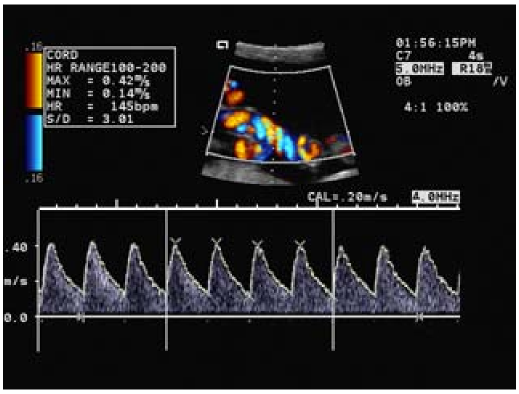
台灣大學電機系李百祺

I. Overview

- The main purpose of all imaging systems is to provide a visual representation of a real object. For medical ultrasonic imaging systems, high frequency sound waves are used to interact with the human body so that anatomical structures, blood flow velocity and other diagnostic information can be obtained.

- A typical ultrasonic imaging system is shown below on the left and an ultrasound image with both anatomical structure of the umbilical cord and associated flow information is shown on the right. (Both are from Acuson corporation, Mountain View, CA.)





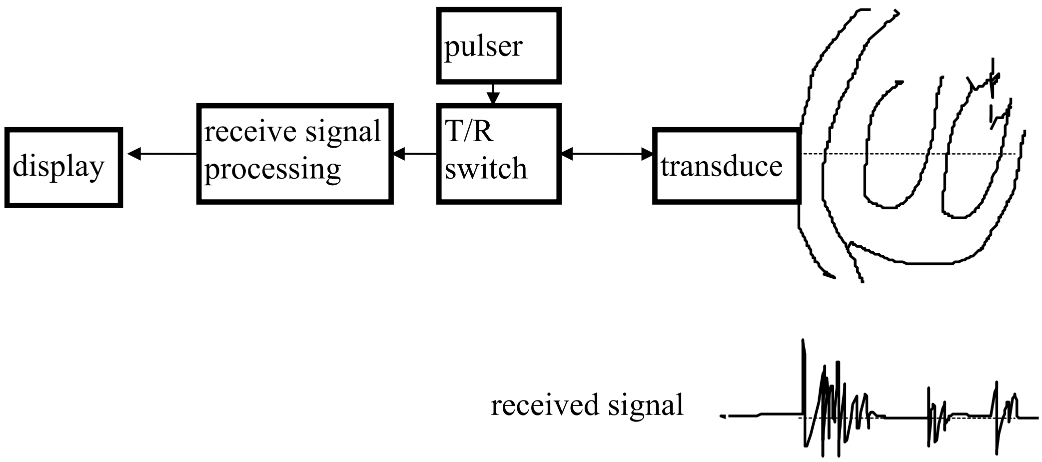
- Ultrasound is the second most widely used diagnostic imaging modality (next to X-ray). Gray scale (B-mode) ultrasound became popular in 1970s. Color Doppler became clinically useful in 1980s. In 1990s, all high-end systems are fully digital. The next generation systems will continue to take advantage of advanced signal processing techniques, faster clock rate, larger memory capacity and higher circuit density.

- The advantages of ultrasonic imaging systems include non-invasive image formation of anatomical structures, non-invasive detection of moving objects, real-time acquisition, portability and cost.

- Current clinical applications include OB/GYN, vascular, cardiac, transcranial, abdominal, musculosckeletal, endo-vaginal, endo-rectal, ocular, intra-vascular, intra-operative...etc. With the recent advancement of electronic technologies and more understanding of the imaging mechanism, the clinical capabilities have been expanding rapidly. Generally, new features are being released to the market approximately every six months.

- Characteristics of ultrasonic imaging:
  - Non-invasive.
  - Safe (under regulations).
  - Direct acquisition of 3D information.
  - Reflection mode, similar to RADAR.
  - Real-time (frame rate is high enough to allow direct interaction between the user and the patient).
  - Velocity estimation of moving targets (such as blood) using Doppler. Information can be displayed either by video or audio.
  - Relatively cost-effective.
  - Access (cannot effectively penetrate air pockets and bone).
  - Resolution determined by diffraction (as in Optics).
  - Modest resolution, but continues to improve.
  - Image quality strongly dependent on "body type".

- A very simple model of ultrasonic imaging (A-scan, one-dimensional)



- A transducer is used to convert an electrical signal to an ultrasonic pulse on transmit, and vice versa on receive (acting as an antenna). The following is a picture of a phased array transducer (Acuson corporation, Mountain View, California).

- The ultrasonic pulse propagates into the body and reflected by tissue inhomogeneities. The reflection is caused by the differences in acoustic impedance which is in turns determined by the propagation velocity and density. Therefore, the returned signal (echo) carries information of the mechanical properties of the image object. The echo is then received by the same transducer and converted into an electrical signal. Due to the round-trip nature of acoustic propagation, the total propagation time for an echo reflected from a depth of R is $t_{roudtrip}=2*R/c$, where c is sound velocity. The propagation time for a typical imaging condition is at the order of a few hundred micro-seconds.

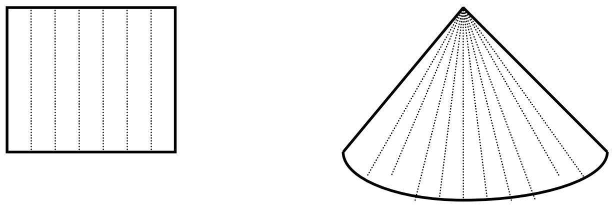
- B-mode (Brightness mode): A two-dimensional image consists of many one-dimensional A-scans (described previously). In other words, a two-dimensional image is acquired by simply moving the directions of the ultrasound beam (i.e., scanning). This is similar to a radar system. The scanning can be done manually, mechanically or electronically. A full view two-dimensional image typically consists of 100-200 lines. Therefore, it is possible to acquire 30 image frames in a second for continuous (real-time) display.

- Current digital systems all use arrays transducers to perform electronic scanning. Such transducers typically consist of 64 to 256 individual transducer elements. The process of using arrays to direct acoustic beams to a particular point in space is called beam formation. Array transducers allow dynamic beam formation (i.e., independent control of individual transducer element) which is not possible with single crystal transducers. However, system complexity, size and cost also dramatically increase.

- Doppler effects have also been used to acoustically detect moving objects in the body (e.g., blood and cardiac muscle). The motion information can be presented by audio speakers, or displayed as Doppler spectra and two-dimensional color images.

- Commonly seen scan formats:

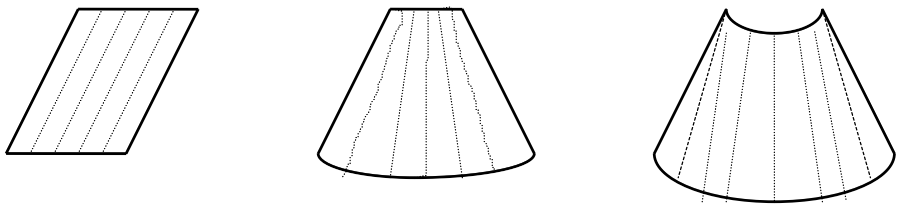
linear | **sector** | **curved linear**



easy access / **limited field of view** | **limited access** / **large field of view** | **easy access** / **large field of view**

- Scan format variations:

steered linear | **vector** | **curved vector**

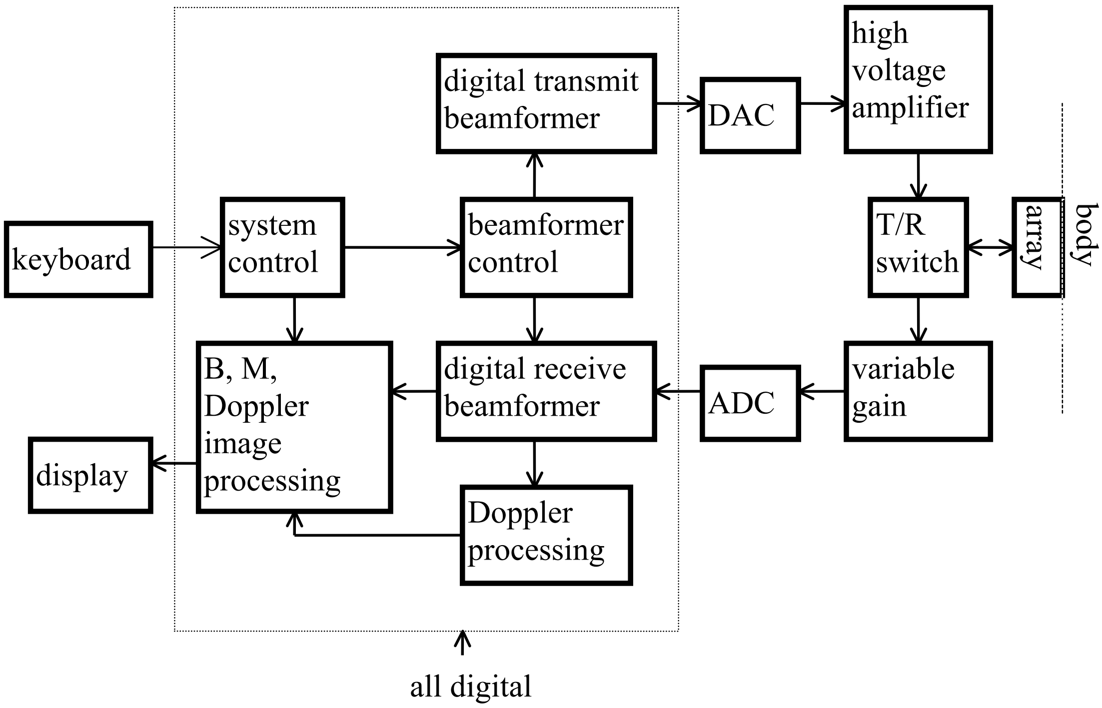


These variations can be used to extend the field of view or to obtain a better angle for flow estimation using Doppler.

- There are two types of three-dimensional ultrasonic imaging. One is to construct three-dimensional object information by using a set of two-dimensional images. The three-dimensional data is then projected to a pre-specified two-dimensional plane for display. This is not real-time.

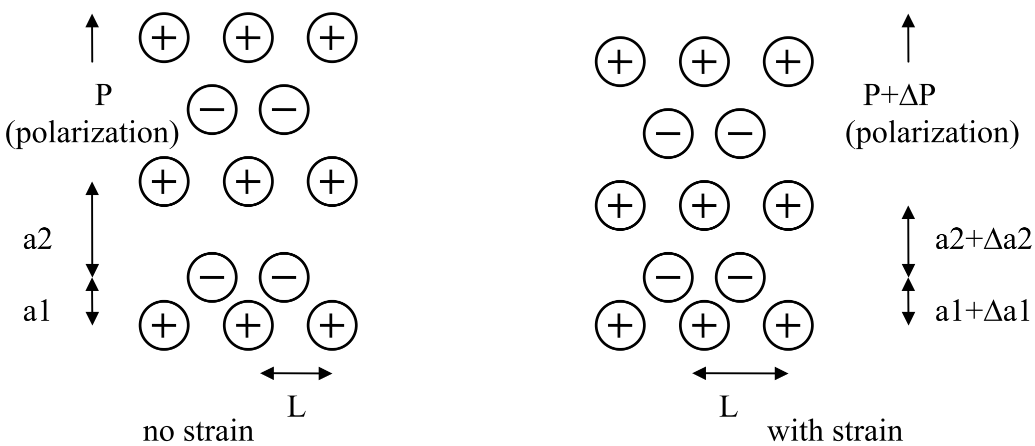
- A real-time three-dimensional ultrasonic imaging system using two-dimensional arrays has recently been commercialized with limited imaging capabilities. Its clinical potential is yet to be realized.

- A fully digital system is a special purpose computer with powerful signal processing capacity, plus transducer arrays and analog front-end. A block diagram of such a system is shown as the following.

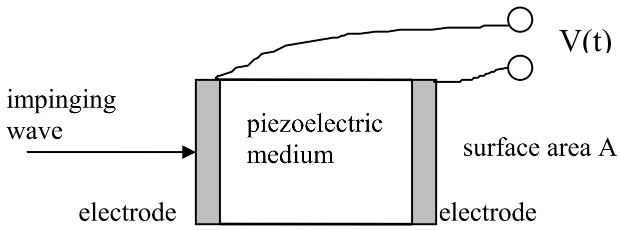


II. Generation and Detection of Ultrasound

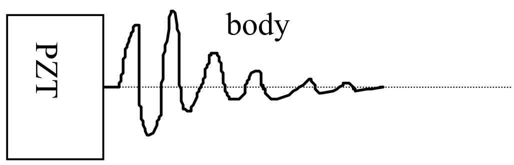
- In diagnostic ultrasound, piezoelectric materials, such as lead zirconate-titanates (PZTs), are most commonly used to generate and detect sound waves. Piezoelectricity is defined as the generation of an electrical polarization in a substance by the application of a mechanical stress and, conversely, a change in the shape of a substance when an electric field is applied. In other words, a material is strained when an electric field is applied to it. Usually the surfaces of a piezoelectric material offer the sharpest discontinuity in the electric field, hence they are the strongest sources of sound.



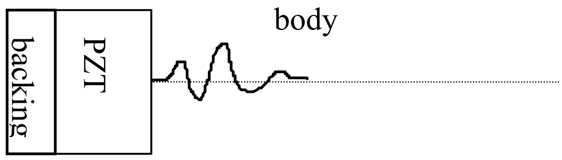
- Piezoelectric detection of ultrasonic waves is reciprocal to the process of wave generation. In other words, the conversion of mechanical energy into electrical energy is also a phenomenon dominated by the behavior at the surfaces of the piezoelectric material.



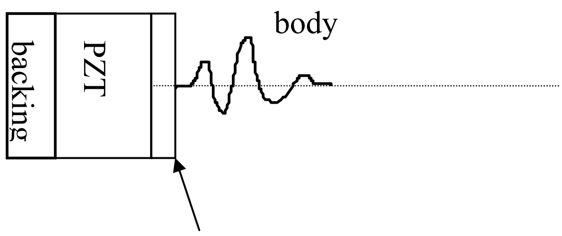
- Broadband transducers are necessary for pulse-echo imaging applications in order to achieve high range resolution, which is inversely proportional to the pulse bandwidth. However, a short pulse is usually achieved by sacrificing sensitivity. Considering the following piezoelectric transducer with no matching or damping layers, it rings and produces an unacceptable long pulse.



By placing a lossy material (highly attenuating) which has a similar acoustic impedance as the PZT, the reflection at the back of the transducer can be reduced and therefore the pulse can be shortened. Apparently, sensitivity is degraded due to attenuation.



The acoustic impedance of a typical PZT material is around twenty times higher than that in the body (30Mrayl vs. 1.5Mrayl), therefore, part of the sensitivity loss can be recovered by adding one or multiple impedance matching layers at the front surface of the transducer.

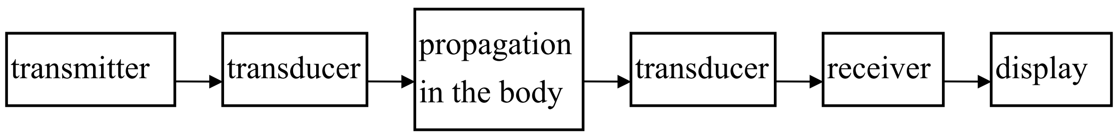


- *Two-way insertion loss,* defined as the ratio of the available electrical power generated by the device as a receiver to the electrical power dissipated in the device as a transmitter under the conditions in which the acoustic wave produced is reflected from a perfectly reflecting interface and received by the same transducer, is often used as a measure of the electromechanical efficiency of the transducer. The lower the insertion loss is, the higher the sensitivity can be achieved.

- An acoustic lens is often placed on the front of the transducer in order to provide a fixed geometric focusing. This is particularly important for imaging using one-dimensional arrays, in which case the geometric focusing is provided along the non-scan direction.

III. Imaging Equations and Diffraction

- Ultrasonic image formation can be described by the following model



Considering one dimensional situations (A-scan), the received signal $V(t)$ is given by

$$V(t) = k\iiint \frac{R(x',y',z')e^{-2\beta z'}}{z'} B(x',y',z')\,p(t-\frac{2z'}{c})\,dx'dy'dz'$$

where $R(x',y',z')$ is the reflectivity of the body at arbitrary position (x',y',z'), $B(x',y',z')$ is the pulse-echo radiation pattern and $p(t-2z'/c)$ is the received pulse-echo signal from an ideal reflector at depth z'.

By scanning in the x direction and assuming the attenuation $(e^{-2\beta z'})$ and the spreading term $(1/z')$ are corrected, the following equation yields

$$S(x,t)=k\iiint R(x',y',z')B(x'-x,y',z')\,p(t-\frac{2z'}{c})\,dx'dy'dz'.$$

Note that without loss of generality, $B(x'-x,y',z')$ can be re-written as $B(x-x',y',z')$ (a simple sign change) such that the above equation can be viewed as a normal convolution in x.

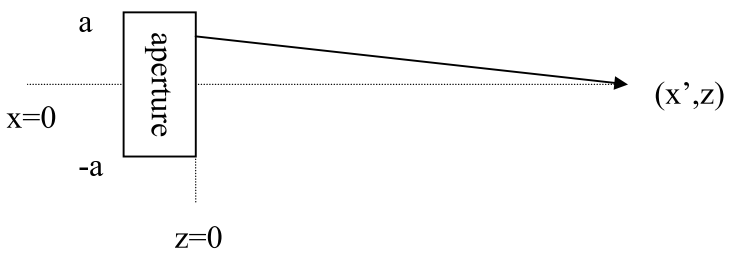
- In general, the pulse-echo waveform can be described by

$$p(t-\frac{2z'}{c})=A(t-\frac{2z'}{c})\cos(2\pi f_0(t-\frac{2z'}{c}))$$

where $A(\cdot)$ is the envelope of the pulse and f_0 is the carrier frequency of the pulse. The width of $A(\cdot)$ determines the axial resolution (resolution in the z direction).

- As in optics, $B(\cdot)$ is mainly determined by diffraction and therefore determines the resolution in a plane ($x-y$ plane) perpendicular to the direction of wave propagation (z direction). For two-dimensional ultrasonic imaging, one-dimensional apertures are adequate.

- The continuous wave (CW) radiation pattern from a finite aperture is simply a superposition of all point sources. Therefore, for a point at (x',z), the pressure wave is



$$p(x',z)=\int_{-a}^{a}\frac{e^{jkd(x,x')}}{d(x,x')}dx$$

where $k=2\pi/\lambda$, λ is the wavelength, $d(x,x')$ is the distance from a point x in the aperture plane to a point x' in plane z and the aperture is non-zero between *(-a, a)*.

- In the Fresnel region, which is defined as $z^2 >> (x-x')^2$,

$$d(x,x')=z(1+\frac{(x-x')^2}{z^2})^{1/2}\approx z+\frac{(x-x')^2}{2z}.$$

and the pressure wave can be approximated by

$$p(x',z)\approx\frac{1}{z}\int_{-a}^{a}e^{jkz}e^{jk(x-x')^2/2z}dx=\frac{e^{jkz}e^{jkx'^2/2z}}{z}\int_{-a}^{a}e^{-jkxx'/z}e^{jkx^2/2z}dx$$

The above equation can be generalized for the case of an arbitrary complex aperture function $C(x)=|C(x)|e^{j\theta(x)}$,

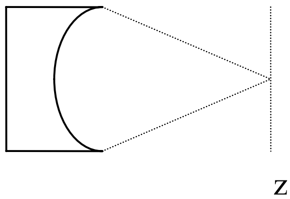
$$p(x',z)\approx\frac{e^{jkz}e^{jkx'^2/2z}}{z}\int_{-a}^{a}C(x)e^{-jkxx'/z}e^{jkx^2/2z}dx$$

- In the far field, where $ka^2/2z<<1$, the integral reduces to

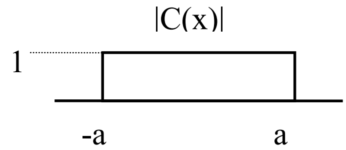
$$p(x',z)\approx\frac{e^{jkz}e^{jkx'^2/2z}}{z}\int_{-a}^{a}C(x)e^{-jkxx'/z}dx$$
$$=\frac{e^{jkz}}{z}F.T.[C(x)]$$

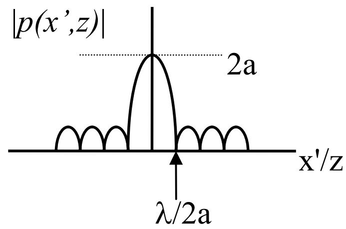
where F.T. stands for Fourier transform.

- When not in the far field, $C(x)e^{jkx^2/2z}$ can be viewed as the effective aperture function. Furthermore, if $C(x)$ is chosen to be $C(x)=|C(x)|e^{-jkx^2/2z}$, the integral reduces to the same equation as in the far field, and hence the Fourier transform relation still holds. An effective aperture $C(x)=|C(x)|e^{-jkx^2/2z}$ simply means a lens focused at a depth z, i.e.,



- For a uniformly weighted aperture, the radiation pattern (in the far field or in focus) is simply a sinc function in combination with some phase terms. In addition, since the width of the diffraction pattern is directly related to the lateral resolution (x' direction), it is clear that the aperture size ($2a$) is inversely proportional to the beam width and therefore proportional to the lateral resolution (i.e., the larger the aperture, the better the lateral resolution). Similarly, the higher the frequency is, the better the lateral resolution can be achieved.





- The above equations are derived for a transmitted wave, the diffraction analysis is exactly the same for reception due to the reciprocal nature. Therefore, assuming a fixed aperture, the two-way radiation pattern becomes

$$B(x',z) = T(x',z,\omega_0)R(x',z,\omega_0) = p(x',z)^2$$

where $\omega_0 = 2\pi f_0$ and f_0 is the carrier frequency. In general, the transmit and receive apertures are different, thus resulting in different radiation patterns.

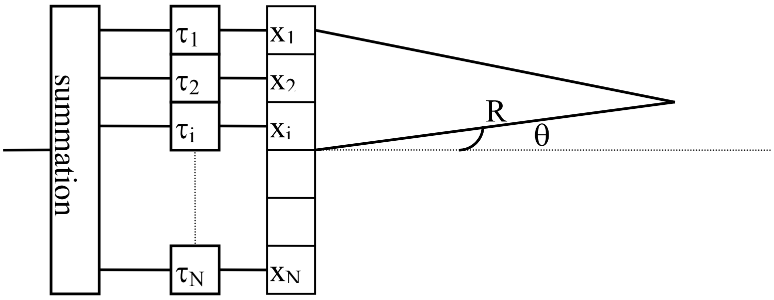
- The actual pulse-echo system uses a pulse (i.e., a broad band signal) rather than a continuous wave. Based on the principle of superposition, the broad band radiation pattern is

$$B(x',z) = \int T(x',z,\omega)R(x',z,\omega)A(\omega)d\omega$$

where $A(\omega)$ is the spectrum of the pulse excitation.

IV. Beam Formation Using Arrays

- In the previous section, we have shown that a lens can be formed by time delaying (i.e., phasing) the aperture. It also becomes apparent how an array can be used to form a beam focused at any particular point in space. The receive beam formation can be illustrated by the following drawing



- Since an array can focus beams at an off-axis point, it can replace the need for mechanical scanning and real-time images can be acquired through electronic focusing and steering.

- By using a sector scan format, any image point in space can be specified by polar coordinates (R,θ). Assuming $S_i(t)$ is the received signal on the i-th element, the diffraction integral shown in the previous section can be approximated by the following coherent summation

$$O(t) = \sum_{i=1}^{N} S_i(t - \tau(x_i, R, \theta))$$

where $O(t)$ is the output signal and

$$\tau(x_i, R, \theta) = \frac{\left((x_i - R\sin\theta)^2 + R^2\cos^2\theta\right)^{1/2}}{c} = \frac{R}{c}\left(1 + \frac{x_i^2}{R^2} - \frac{2x_i}{R}\sin\theta\right)^{1/2}.$$

In the Fresnel region,

$$\tau(x_i, R, \theta) \approx \frac{R}{c}\left(1 + \frac{x_i^2}{2R^2} - \frac{x_i}{R}\sin\theta - \frac{x_i^2}{2R^2}\sin^2\theta\right)$$

$$= \frac{R}{c}\left(1 - \frac{x_i}{R}\sin\theta + \frac{x_i^2}{2R^2}\cos^2\theta\right) = \frac{R}{c} - \frac{x_i\sin\theta}{c} + \frac{x_i^2\cos^2\theta}{2Rc}$$

- The first term in the above equation is independent of the angle and the channel index, it is simply the propagation time from the center of the array to a range *R*. The second term relates to the beam direction and is independent of the range. It represents the steering component. The third term is a parabolic function and represents the focusing component. It approaches to zero in the far field.

- The third term in the above equation also implies that when the focal point is off-axis ($\theta \neq 0$), the effective aperture size also reduces from $2a$ to $2a\cos\theta$. In other words, the lateral resolution is reduced when the scan angle is steered off normal.

- On transmit, since the array is only fired once, a fixed focus must be specified for the transmit beam (for that particulr firing). On receive, on the other hand, since the received signal is continuously stored as a function of time, the receive focus can be continuously updated. Therefore, the receive signal can be dynamically focused in the image range.

- As in any other sampling problems, the choices of the interelement spacing (i.e., pitch) in an array and the beam sampling interval for a two-dimensional image need to satisfy Nyquist criteria in order to avoid aliasing. The detailed analysis, however, is beyond the scope of this course.

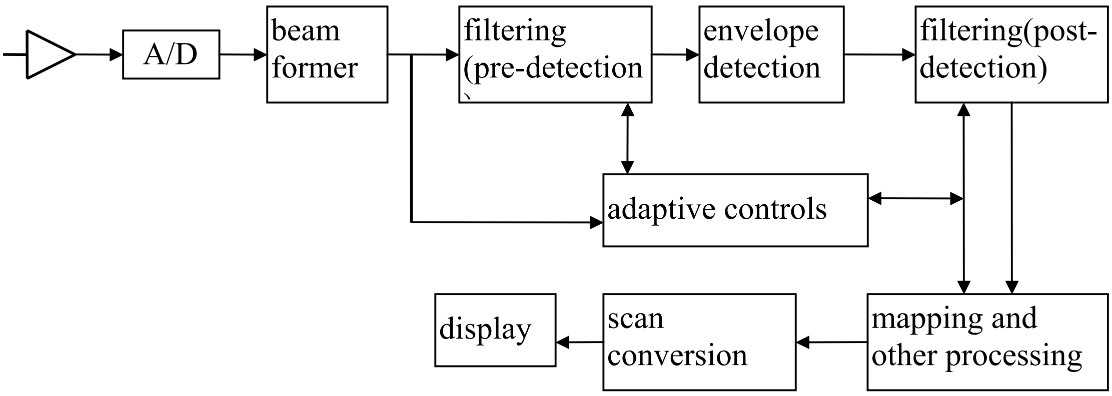
.

V. Real-Time Image Formation

- An ideal transmitter is capable of sending out a pulse with an arbitrary waveform (i.e., arbitrary pulse shape, amplitude and time delay). The pulse shape determines the frequency contents of the transmitted signal. The amplitude of the pulse varies across the array in order to provide aperture weighting (a.k.a. apodization). It can also be varied in order to adjust the acoustic power delivered to the body. Finally, the time delay for each individual channel is related to focusing quality.

- Since a fixed transmit focus is usually used for each firing, the transmitter hardware is relatively easy. In addition, some systems are only using an impulse for the pulse excitation such that system cost can be further reduced.

- A generic digital receiver is illustrated in the following figure



- Signals received by the transducer array are first amplified in order to compensate for the depth dependent signal loss during propagation. This is also known as time gain compensation (TGC). The amplified signals are then digitized and sent to the receive beam former for dynamic receive focusing.

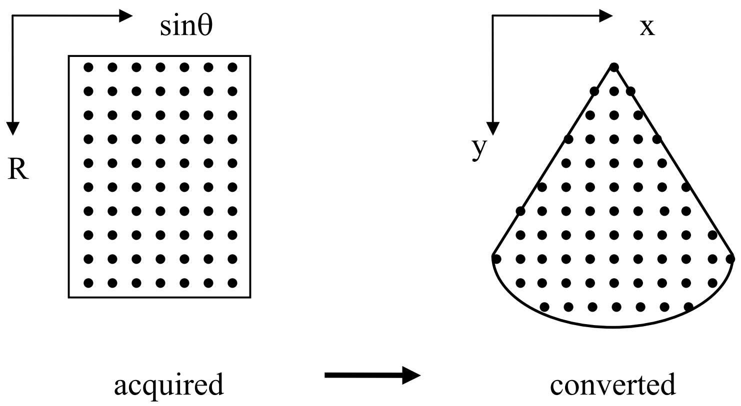
- The pre-detection filter following the receive beam former serves the purpose of pulse shaping (axial filtering), beam shaping (lateral filtering) and temporal filtering (filtering between image frames). In other words, it is used to alter the frequency contents (both temporal and spatial) of the received signals. Additionally, it can be used for data re-sampling (interpolation or decimation).

- Typically, only the envelope of the received signal is displayed. The envelope detector is used to filter out the carrier frequency component and to keep only the magnitude of the remaining signal (e.g., $|A(t)|$). The envelope detected signal is also called the video signal.

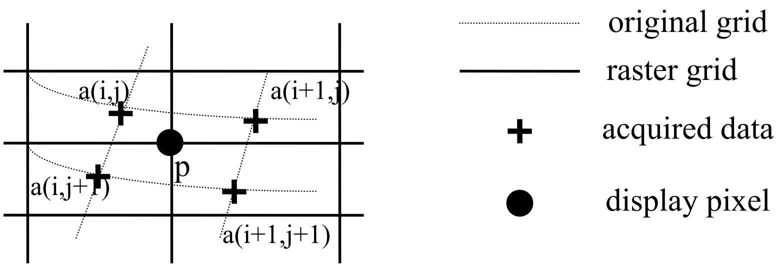
- The post-detection filter is applied to the detected signal for the purpose of smoothing, edge enhancement or a combination of both. It is typically used to enhance contrast or spatial resolution such that particular features in the image can be better perceived. Additional data re-sampling can also be performed at this stage. Other processing, such as re-mapping (gray scale or color) and digital gain adjustment, can be applied following the post-detection filter.

- Ideally, the imaging parameters in all the aforementioned functional blocks should be controlled and adjusted adaptively. In other words, the system should be optimized automatically based on clinical needs, patient characteristics and users' personal preference. However, due to implementation difficulties, most systems are not fully adaptive and the optimization is usually done manually.

- Scan conversion is the process of converting the acquired ultrasound data (e.g., polar coordinate data for a sector scan format) into the Cartesian raster data used by standard display monitors. In order not to introduce artificial artifacts, such as a Moire pattern, two-dimensional interpolation is usually involved for most pixels in the image and hence it is computationally demanding. By using sector scan format as an example, an image with a sector scan format is acquired on a polar grid (R,θ) and needs to be converted to a Cartesian raster grid (x,y) for display.



- Generally the exact location of an acquired data point does not correspond to an allowable display pixel. Hence, the data must be converted before in can be written to the display buffer. This conversion is typically done by bi-linearly interpolating the nearest acquired data points shown below.



Let $a(i,j)$ be the original acquired sample at the i-th line and the j-th range point. If the output pixel $p(m,n)$ (the (m,n) point on the raster grid) falls in between $a(i,j), a(i+1,j), a(i,j+1)$ and $a(i+1,j+1)$, then

$$p(m,n) = c_{m,n,i,j}a(i,j) + c_{m,n,i+1,j}a(i+1,j) + c_{m,n,i,j+1}a(i,j+1) + c_{m,n,i+1,j+1}a(i+1,j+1),$$

where $c_{m,n,i,j}$'s are pre-determined interpolation coefficients and are a function of the distance between $p(m,n)$ and four neighboring points.

- With less accurate interpolation schemes (e.g., nearest neighbor mapping), image artifacts, such as the Moire pattern, may be present. The artifacts are not only aesthetically unpleasant, they may also interfere with the true diagnostic information in the image.

VI. Temporal Resolution

- Fundamentally, frame rate (i.e., number of image frames per second) is determined by the propagation time of sound waves. For a sound velocity of 1.54*mm/usec* in soft tissue, the round trip propagation time required for a depth of 20cm is simply $t_{line} = 2 \cdot 200 / 1.54 \approx 259.74 u\sec$. On the other hand, a typical image consists of around 100 to 150 ultrasound lines (determined by the spatial sampling criterion). Therefore, the frame time required for an image consisting of 120 lines is $t_{frame} = 120 * t_{line} \approx 31.17 ms$ and the frame rate is $1/31.17 ms \approx 32 frames/\sec$.

- Generally speaking, images are displayed in "real-time" if the frame rate is around or above 30 (30 for NTSC standards and 24 for motion pictures with 2:1 interlace). For a typical B-mode image, such as the above example, the frame rate is high enough such that the perception of flickering is at its minimum. However, there are many situations in which the frame rate is significantly lower than 30Hz (e.g, 10Hz) due to dense line sampling and/or repetitive firing for blood flow estimation. Under these situations, the motion of tissue may not appear very smoothly. However, the frame rate is still high enough such that users can easily interact with patients and adjust the image in a timely fashion. This type of direct interaction between the users and the patients is indeed the essence of "real-time" diagnostic imaging.

- It is also likely that the frame rate is much higher than 30Hz (e.g., 100Hz).

Although information may be lost during live scanning for systems which can only refresh the screen at 30Hz, the information can still be used by displaying still images or by performing slow motion video playback. This is why high frame rates (higher than 30Hz) are still diagnostically important for fast moving structures (such as heart valves) .

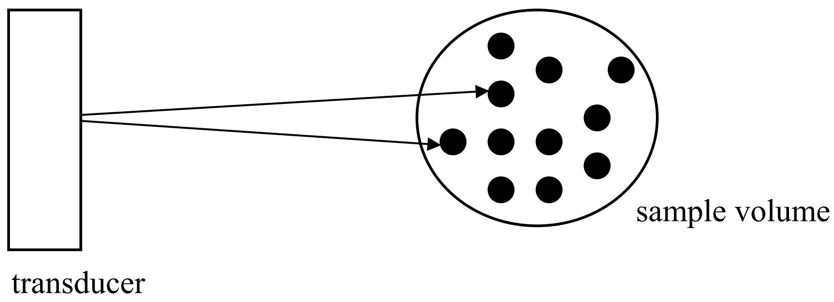
- Although a 30Hz frame rate with 2:1 interlace is sufficient to avoid the perception of flickering, it does not guarantee that the temporal sampling rate satisfies the Nyquist criterion. This is also why higher frame rates may be clinically useful and an important feature of a system. There are many methods that can increase the frame rate (e.g., by parallel beam formation or by increasing the beam sampling interval). The details, however, are beyond the scope of this course.

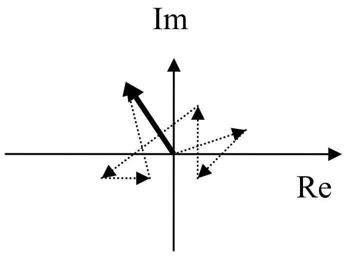
VII. Speckle and Contrast Resolution

- Speckle is a common phenomenon in coherent (i.e., phase sensitive) imaging systems. It comes from coherent interference of scatterers (i.e., linear summation by the transducer of echoes from scatterers) and it appears as a granular structure superimposed on the image. Speckle is an artifact degrading target visibility and does not represent any inherent tissue properties. It is one of the primary limitations on detecting low contrast lesions in ultrasonic imaging.

- In ultrasound, speckle occurs when the size of scatterers is small compared to a wavelength and there are many scatterers within one sample volume. This is true for diagnostic ultrasound when imaging soft tissue at typical frequencies. In this case, the speckle pattern statistics are independent of the scattering structures and are a function only of the imaging system and its relative distance to the target.

- Speckle can be modeled by a random walk in the complex plane where each step in the walk represents the signal received by the transducer from a single scatterer within the resolution volume. Since these scatterers are in the same sample volume, the signals from them are coherently summed when received by the system. The sum can therefore be represented by a complex number shown in the following drawing. Note that the amplitude of each individual phasor represents the strength of the signal from the particular scatterer, whereas the phase of the phasor is related to the propagation delay (i.e., distance between the scatterer and the transducer).





- Let a_k $(1 \le k \le N)$ be the signal received from a particular scatterer with a phase θ_k ($\theta_k = 2\pi R/\lambda$, where R is the distance between the scatterer and λ is the wavelength), we have the following equations for the normalized summed signal A

$$A = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} |a_k| e^{i\theta_k}$$

Since N is large, the central limit theorem is applicable. Defining

$$\sigma^2 = \frac{1}{N} \sum_{k=1}^{N} \frac{|a_k|^2}{2}$$

$$I \equiv |A|^2$$

It can be shown that the probability density function of the intensity of the summed signal I ($I \ge 0$) is as the following (exponential distribution)

$$p_I = \frac{1}{2\sigma^2} e^{-\frac{I}{2\sigma^2}}$$

The amplitude $E \equiv \sqrt{I}$ ($E \ge 0$) has the following distribution (Rayleigh distribution)

$$p_E = \frac{E}{\sigma^2} e^{-\frac{E^2}{2\sigma^2}}$$

- Some interesting properties for the above two density functions are listed as follows:

$$SNR_I \equiv \frac{\langle I \rangle}{\sigma_I} = 1$$

$$SNR_E \equiv \frac{\langle E \rangle}{\sigma_E} = \frac{\left(\pi\sigma^2/2\right)^{1/2}}{\left(\left(4-\pi\right)\sigma^2/2\right)^{1/2}} \approx 1.91$$

where σ_I and σ_E are standard deviations of the intensity and the amplitude, respectively.

Therefore, speckle noise can be viewed as a multiplicative noise, where the noise increases as the mean increases. In other words, stronger signals also suffer from higher noise and hence they are not easier to be detected.

- Ultrasonic images are often displayed on a logarithmic scale so that signals across a wide dynamic range can be shown at the same time. In this case, the speckle noise becomes *additive* on a logarithmic scale. Define *D* as

$$D(dB) = f(I) \equiv 10\log_{10}(\frac{I}{I_0})$$

where I_0 is an arbitrary reference signal. Expanding the function *f* in a Taylor series about the statistical mean $\langle I \rangle$ and keep only the zero-th and the first order terms, we have

$$\sigma_D{}^2 \approx f'(\langle I \rangle)^2 \sigma_I{}^2 = \left(\frac{10}{\ln 10}\right)^2 \frac{\sigma_I{}^2}{\langle I \rangle^2}$$

$$\sigma_D \approx 4.34(dB)$$

In other words, speckle noise on a logarithmically processed display becomes a fixed additive noise. This noise (4.34dB) fundamentally limits the contrast resolution in diagnostic ultrasound.

- Contrast resolution, defined as the ability to detect objects with different sizes and mean brightness against the background, is an important feature of an ultrasound system. In diagnostic ultrasound, examples of detecting such objects include the detection of tumors in the liver, myocardium infarction and breast masses.

- Detectability of a lesion in a noisy background is determined by local contrast and the properties of the system. Local contrast, defined as the ratio of the intensity variation to the averaged background intensity, can be artificially enhanced, e.g., by re-mapping gray levels of the envelope. Noise characteristics, on the other hand, inherently limit detectability. To detect a finite lesion, statistics of averaged signals over the regions of interest (i.e., area-wise statistics instead of point-wise statistics) must be derived. More specifically, a contrast-to-noise ratio (CNR), defined as the ratio of the local intensity variation to the standard deviation of the averaged background intensity, is suitable for a quantitative measure of detectability.

- If *I(x',z)* denotes the intensity at a point *(x',z)* in the image plane, the averaged signal over a target is

$$I_A = \frac{1}{S}\iint_{-\infty}^{\infty} W(x',z)\,I(x',z)\,dx'dz$$

where $W(x',z)$ is a weighted function describing the target and

$$S = \iint_{-\infty}^{\infty} W(x',z)\,dx'dz\,.$$

In our analysis, $W(x',z)$ is chosen to be unity inside and zero outside so that the constant S simply represents the target area. Assuming the **point-wise** first order moment (i.e., statistical mean) of $I(x',z)$ in the target region is a constant $\langle I \rangle$, i.e., the target has a homogeneous distribution of scatterers, the mean value of the **area-wise** averaged signal I_A is also $\langle I \rangle$ according to the principles of superposition.

- The second order moment of I_A is

$$\langle I_A^2 \rangle = \frac{1}{S^2}\iiint\int_{-\infty}^{\infty} W(x_1',z_1)W(x_2',z_2)\langle I(x_1',z_1)\,I(x_2',z_2)\rangle dx_1'dx_2'dz_1dz_2\,,$$

where $\langle I(x_1',z_1)\,I(x_2',z_2)\rangle$ represents the autocorrelation function of the intensity. Assuming the speckle pattern is spatially stationary and assuming the target area is much larger than the area where $\langle I(x_1',z_1)\,I(x_2',z_2)\rangle$ is appreciable, we have

$$\sigma_{I_A}^2 = \frac{1}{S}\iint_{=\infty}^{\infty} C_I(x',z)\,dx'dz$$

where $C_I(x',z)$ is the autocovariance function of $I(x',z)$ and σ_{I_A} is the standard deviation of I_A (i.e., $\sigma_{I_A}^2 = \langle I_A^2 \rangle - \langle I_A \rangle^2$).

- Defining the normalized correlation cell area S_c of the target as

$$S_c \equiv \iint_{-\infty}^{\infty} \frac{C_I(x',z)}{C_I(0,0)} dx'dz$$

the standard deviation of the averaged signal, I_A, is

$$\sigma_{I_A} = \frac{\sigma_I}{N^{1/2}},$$

where $\sigma_I^2 = C(0,0)$ is the variance of $I(x',z)$ and N represents the number of independent speckle spots in the target region defined by

$$N = \frac{S}{S_c}.$$

- From the above analysis, it is clear that the standard deviation of the area-wise signal (i.e., I_A) is determined by both the standard deviation of the point-wise signal (i.e., I) and the number of speckle correlation cells. Therefore, defining ΔI as the local intensity variation (i.e., $\Delta I \equiv I_1 - I_2$), then the CNR is

$$CNR = \frac{\langle \Delta I \rangle}{\sigma_{I_A}} = \frac{\langle \Delta I \rangle}{\sigma_I} N^{1/2}$$

where I_1 and I_2 are mean intensities of the target and the background respectively.

- As discussed previously, intensities of fully developed speckle have exponential statistics and speckle fluctuations can be viewed as additive noise on a logarithmic display. Furthermore, the brightness variation between the target and the background on a logarithmic display is

$$\Delta D = 10\log_{10}(\frac{I_1}{I_0}) - 10\log_{10}(\frac{I_2}{I_0}) = 10\log_{10}(\frac{I_1}{I_2}).$$

Using σ_D to denote the standard deviation of the display intensity on a logarithmic scale, we have

$$CNR = \frac{\Delta D}{\sigma_D} N^{1/2} = \frac{10\log_{10}(\frac{I_1}{I_2})}{\sigma_D} N^{1/2}.$$

Finally, recall that the speckle variation on a logarithmic scale is approximately 4.34dB, we obtain

$$CNR = \frac{10\log_{10}(\frac{I_1}{I_2})}{4.34} N^{1/2}$$

In other words, the detectability of a lesion in a speckle background is determined by three factors : the ratio of mean intensities between the target and the background, the total number of speckle correlation cells in the target, and speckle variations.

- Contrast measured by the system is fundamentally limited by the differences in reflectivity between the region of interest and the surrounding tissue. In practice, the measured contrast is even lower than the true contrast due to imperfection in beam formation. Sidelobes, resulting from a finite aperture size, receive signals from directions outside of the main beam direction and therefore may decrease the contrast assuming the signal level in the main beam direction is lower than the background. There are other sources of sidelobes, such as beam formation errors, contributing to degradation of contrast resolution via a similar mechanism. Therefore, lower sidelobe levels often produce higher contrast resolution and this explains why non-uniform apodization, instead of uniform apodization, is often preferred.

- Contrast resolution can be enhanced by decreasing speckle variations (i.e., σ_D). Since speckle is due to coherent interference from scatterers within a sample volume, speckle variations can only be reduced by processing **in-coherently** (i.e., discarding phase information). For example, speckle variations can be reduced by low pass filtering the **post-detection** data, since phase information has been discarded at this point and any processing done thereafter is in-coherent. Speckle variations cannot be reduced by low pass filtering the **pre-detection** data since linear combinations of phase sensitive data do not change the statistical properties of the resulting signals. Note that although speckle noise can be reduced by low pass filtering the post-detection data, contrast resolution enhancement is usually gained at the price of spatial resolution and edge definition.

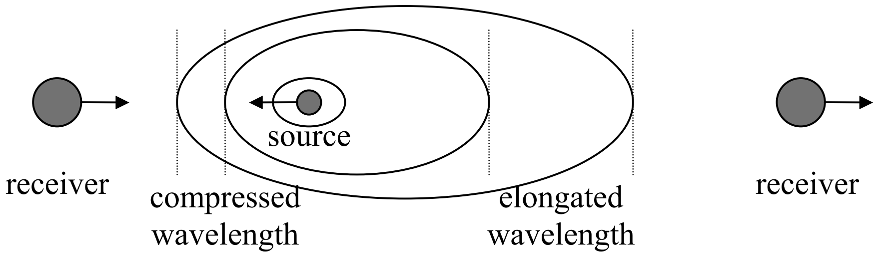
- Compounding techniques (another type of in-coherent averaging) have also been proposed for speckle reduction by introducing in-coherence during image formation. Common compounding techniques include spatial compounding, frequency compounding and temporal compounding. These techniques exploit the fact that the ensemble average of a speckle image is the same as the in-coherent average of the original object. Hence, speckle variations are reduced by in-coherently averaging partially correlated measurements without affecting the original intensity contrast. Generally speaking, partially correlated images can be obtained in spatial (by laterally translating the aperture), frequency (by dividing a pulse into sub-bands) or temporal (by combining pixels that are sampled at a higher rate in time) domain.

- Another way to enhance contrast resolution is to increase the number of independent speckle spots (i.e., correlation cells) in the region of interest. Since the size of the object is fixed, this can only be achieved by decreasing the size of speckle spot. In other words, by improving the spatial resolution, there are more independent samples within the object available to the observer (either the imaging system or the human visual system) such that speckle variations can be smoothed out by additional processing. Given a transducer size and a carrier frequency, it is usually impossible to achieve optimal spatial and contrast resolution simultaneously due to the interaction between mainlobe and sidelobes. Therefore, the only way to improve both spatial and contrast resolution at the same time is to change the fundamental diffraction limits by increasing the aperture size and using a higher carrier frequency.

VIII. Ultrasonic Doppler Imaging

(Doppler Effect)

- In addition to imaging anatomical structures in the body, sound waves can also be used to detect moving objects using the Doppler effect. The most important clinical application of Doppler ultrasound is blood flow measurement. The basic Doppler principles are shown in the following figure.



- When both the source and the receiver are stationary, the frequency experienced by the receiver is simply $f_s = c/\lambda$, where c is the speed of sound and λ is the wavelength. Suppose the source is stationary and the receiver is moving toward the source at a velocity of v_r, the listener experiences $(cT/\lambda + v_r T/\lambda)$ waves in time T. Therefore, the frequency experienced by the receiver is $f = (c + v_r)/\lambda$ and the Doppler shift frequency is $f_d = f - f_s = v_r f_s / c$. Note that generally v_r can also be a negative number representing the receiver moving away from the source.

- When the source is moving and the receiver is stationary, the Doppler effect is different from the above situation. When the source moves toward the stationary listener, the wavelengths are in effect shortened because between emission of two consecutive waves, the source has moved. Suppose the source velocity is v_s, the wavelength is compressed from c/f_s to $(c - v_s)/f_s$. Hence, the frequency experienced by the receiver becomes $f = f_s c/(c - v_s)$ and the Doppler shift frequency is $f_d = f - f_s = f_s v_s /(c - v_s)$. Similar to the previous situation, v_s can also be a negative number indicating the source moving away from the receiver.

- If both the source and the receiver are moving, the Doppler shift frequency becomes

$$f_d = f_s (v_r + v_s)/(c - v_s) .$$

Assuming the source velocity is much smaller the sound velocity (i.e., $v_s << c$), the above equation can be simplified as

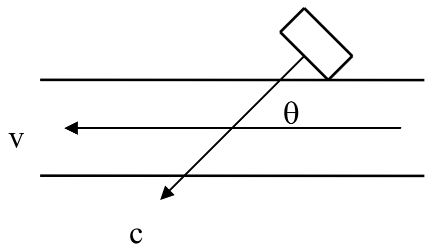
$$f_d \approx f_s (v_r + v_s)/c .$$

This assumption is true for diagnostic ultrasound since the velocity of typical physiological flows in the body are at most at the order of a few meters per second, whereas sound velocity in blood is around 1500 meters per second.

- Since motion of the source relative to the receiver causes a change in the observed sound frequency, blood flow velocity can be measured by detecting Doppler frequency shift of echoes backscattered from moving blood. The primary scattering site in the blood that produces echoes is the red blood cell. The scattering cross section of the red blood cell is about 1000 times larger than that of the platelet. Leukocytes are not present in sufficient numbers to influence the total backscattered signal.

- The incident ultrasonic pulse extends over many red blood cells (there are a few million red blood cells per mm^3 and they are not resolvable by typical ultrasonic imaging systems). Therefore, the signals reflected from the red blood cells within a sample volume are generally not in phase. However, they do vary in unison, i.e., they do change in an orderly fashion such that their velocity can still be estimated by using the Doppler effect.

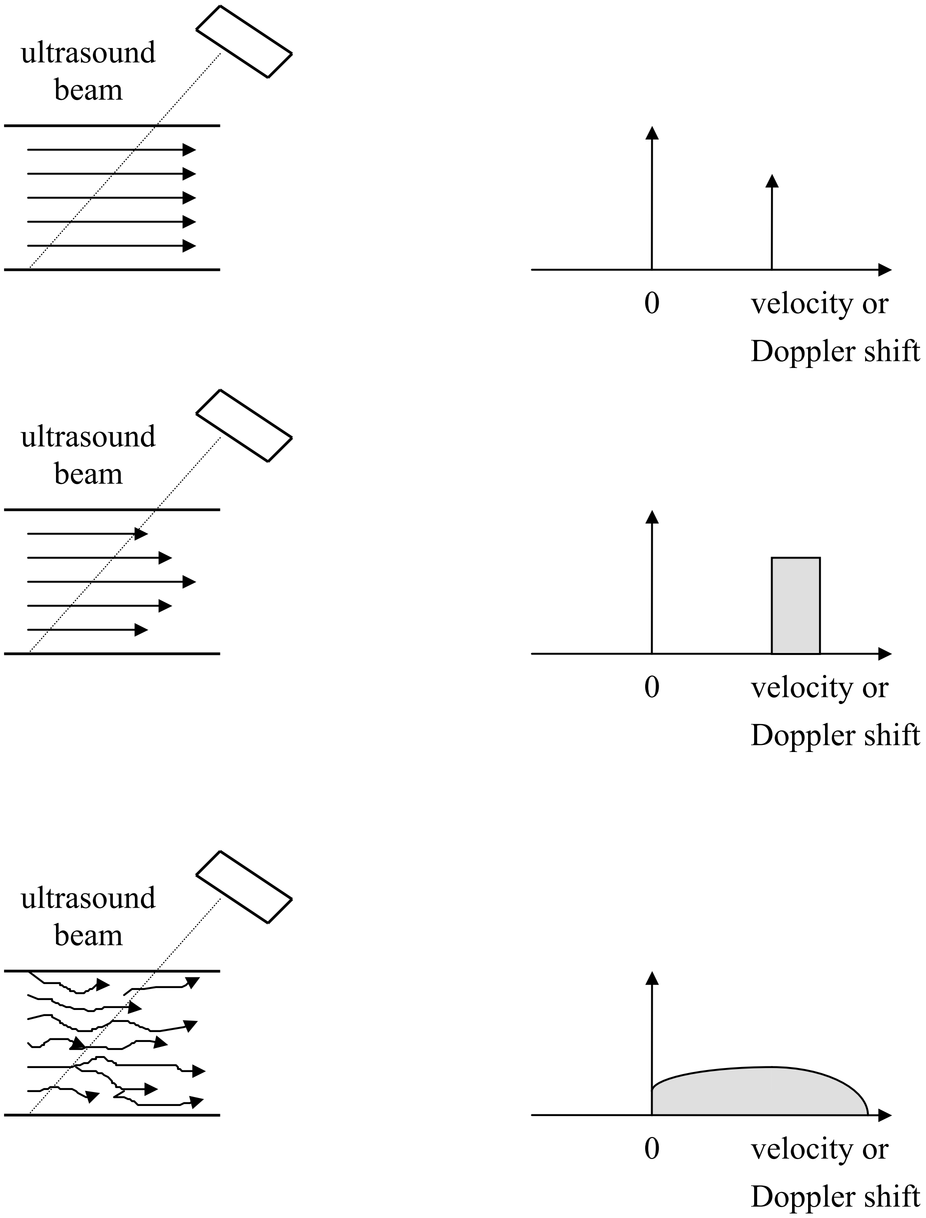
- Considering the following drawing, the basic Doppler equation can be written as



$$f_d = \left(\frac{2vf_s}{c}\right)\cos\theta,\ (v << c),$$

where v is the flow velocity, c is the propagation velocity of sound waves, f_d is the detected Doppler frequency shift, f_s is the source frequency and θ is the angle between the flow and the ultrasound beam. Note that the frequency shift is doubled due to round trip propagation and only flows parallel to the ultrasound beam can be detected.

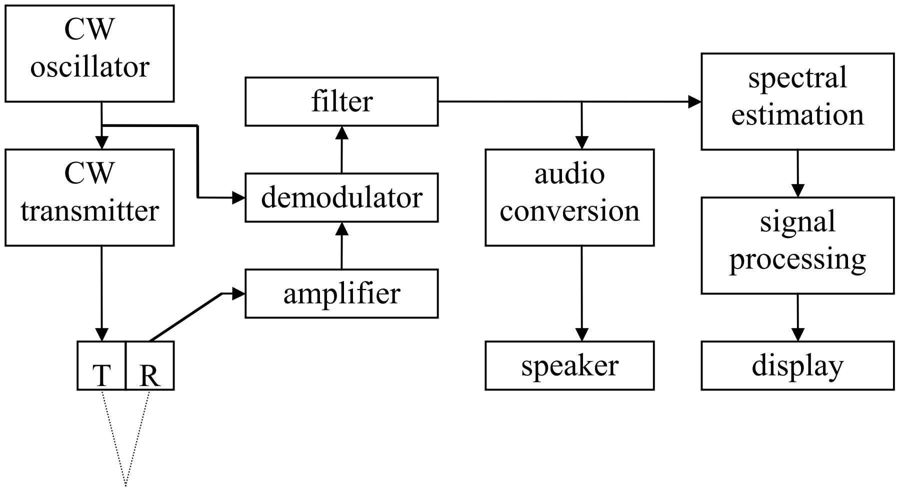
- Flow patterns and their corresponding velocity profiles can be illustrated by the following examples:



- Frequency shifts due to Doppler effects are typically estimated by short-time Fourier transform or efficient auto-correlation techniques. The former is usually called Spectral Doppler, in which only flows along a single ultrasound line (continuous wave, CW) or flows within a particular sample volume (pulsed wave, PW) are estimated. The latter only estimates simple flow parameters, such as mean velocity, flow energy and velocity variance, and is capable of displaying two-dimensional flow information in real-time. Since flow parameters (velocity, variance and energy) are mapped to different colors, this is also known as the Color Doppler mode.

(CW Doppler)

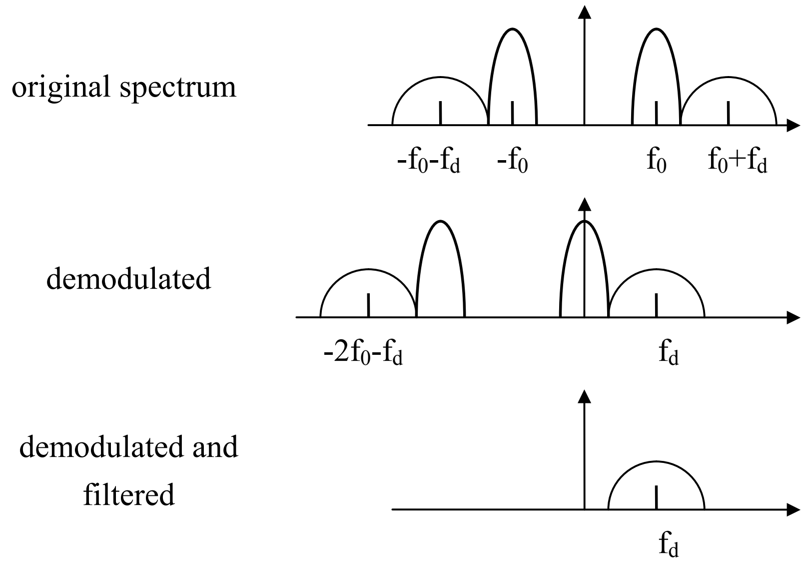
- The first reported application of the Doppler effect in medical ultrasound was CW (continuous wave) velocity measurement of blood. A basic CW Doppler instrument is illustrated in the following drawing.



- In CW mode, one half of the transducer array is continuously transmitting and the other half is continuously receiving. Alternatively, a special transducer made with two piezoelectric crystals for transmitting and receiving the ultrasonic wave can be used. The former is known as array CW and the latter is known as AUX (auxiliary) CW. Note that the AUX CW transducer is only for Doppler measurement and is not suitable for two-dimensional imaging.

- Since one half of the transducer is continuously transmitting, range information cannot be extracted from the returning echoes. In other words, flow information along the entire ultrasound beam is measured by the receiver. Additionally, the frequency downshift due to attenuation can be ignored since the ultrasonic wave is a very narrow band signal (close to zero bandwidth).

- The Doppler shift frequencies are extracted based on the following steps. First, the original signal is demodulated down to baseband. Second, the demodulated signal is low pass filtered to filter out higher frequency components. Third, a high pass filter (a.k.a. wall filter) is used to filter out non-moving and low velocity signals from stationary tissue and relatively slow moving vessel walls. These signals are generally much stronger than the flow signals. The resultant Doppler spectrum is then estimated and displayed. The frequency domain representation of this process can be illustrated by the following drawing. It is worth mentioning that returning echoes from blood are generally highly non-stationary. Therefore, extraction of flow information is indeed a time-frequency analysis.



- The first spectral displays were time-interval histograms made by measuring the times between zero crossings of the Doppler shift waveform. The time intervals are then lumped into bins and displayed as a histogram that gets updated many times in a heart cycle. Modern ultrasonic imaging systems typically use short time Fourier transform. In this case, signal in a moving window (in time) is Fourier transformed and the resultant spectrum represents the velocity distribution of blood flow at that particular instance. It is assumed that the signal within the window of analysis is stationary and the Fourier transform is typically done by 32-128 points FFT (Fast Fourier Transform).

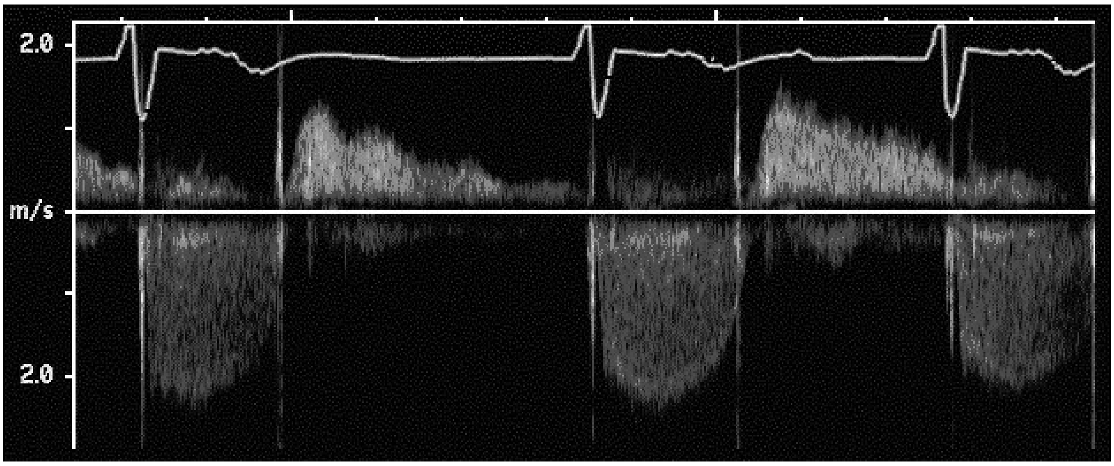
- Other time-frequency analysis techniques, such as AR (auto regressive), wavelet and adaptive cone kernel techniques, have also been explored. However, if only the general features of the flow distribution are of interest, the short time Fourier transform is the most stable and interpretable. Other techniques, in particular model based techniques, tend to break down when the model does not fully characterize the flow.

- Typically only magnitudes of the estimated spectra are displayed and the magnitudes are generally converted to a logarithmic scale. In addition, filtering may be applied in both the frequency direction and the temporal direction prior to display.

- For blood velocities present in the body and for frequencies typically used by the ultrasonic imaging systems, the Doppler shift frequencies from blood happen to occur in the human audible range (near DC to 20KHz). In fact, early CW instruments simply played the Doppler shift frequencies into a speaker. Modern CW instruments differentiate positive frequency shifts from negative frequency shifts. Positive ones, which correspond to flows moving toward the transducer, are played in one channel of a stereo pair of speakers and negative ones, which correspond to flows moving away from the transducer, are played in the other channel. This mode is also known as Audio Doppler.

- Since human ear is very adept at recognizing frequencies and bandwidths of signals in the presence of background white noise, Audio Doppler can offer great diagnostic information once the operators were trained well enough to recognize the normal Doppler shift sounds from abnormal flow patterns. It is not unusual that weak flows are "heard" through Audio Doppler but are hardly "seen" in the CW Doppler strip.

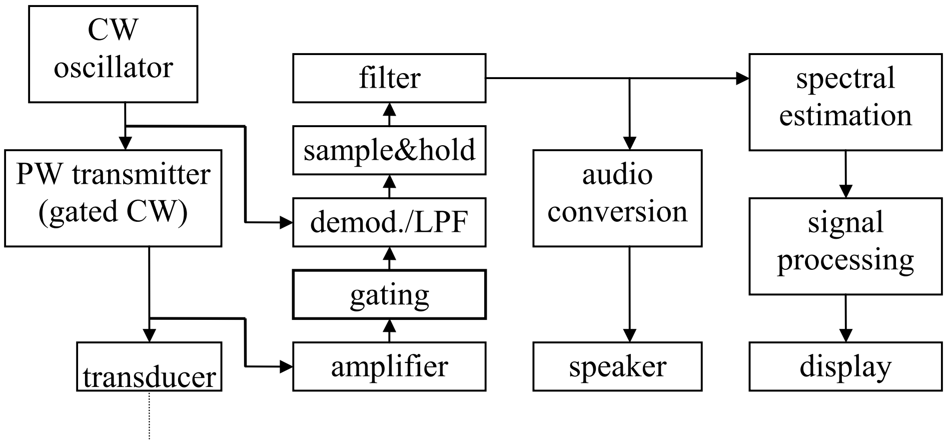
- A typical CW Doppler strip is shown in below.



(PW Doppler)

- The major limitation of CW Doppler is that it is sensitive to blood flows along the entire length of the ultrasonic beam and the position of individual flows cannot be localized in range. To overcome this problem, pulsed Doppler was developed.

- A basic PW Doppler instrument is illustrated in the following drawing.



- One of the main differences between CW Doppler and PW Doppler is that PW Doppler transmits a short sinusoid burst, instead of a continuous sine wave. By gating the received signals to correspond to the pulse's time of flight to the point of interest, one can interrogate a small sample volume instead of the entire length of the ultrasonic beam.

- Unlike CW Doppler, frequency (i.e., velocity) aliasing may occur in PW Doppler. This is due to the fact that the returned echoes are sampled at a fixed time after the burst is transmitted. This time interval is also known as PRI (pulse repetition interval). According to the Nyquist criterion, the maximum Doppler shift frequency ($f_{\max}$) that can be detected for a given PRI is

$$2f_{\max} \le \frac{1}{PRI}.$$

Therefore, the maximum velocity ($v_{\max}$) of the flow projected along the direction of t he ultrasonic beam that can be detected without aliasing is

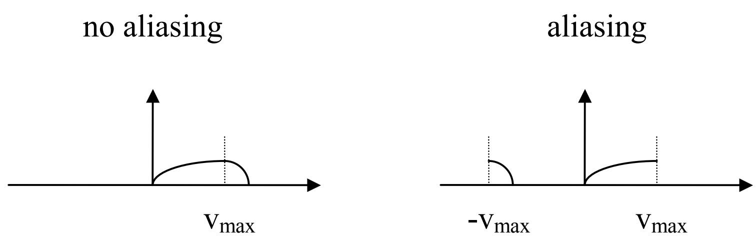
$$2f_{\max} = \frac{4v_{\max} f_s}{c} \le \frac{1}{PRI}$$

$$v_{\max} \le \frac{\lambda}{4 \cdot PRI}$$

where λ is the wavelength at the carrier frequency (i.e., without Doppler shift).

- Due to the sampling requirement, a trade-off exists between the maximum sample depth and maximum un-aliased blood velocity. When aliasing occurs, the high velocity components (higher than $v_{\max}$) show up in the other side of the spectrum.

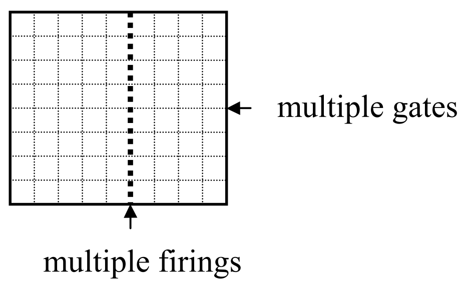
This is illustrated as follows.



- The CW spectral estimation and signal processing techniques are also applicable for PW. Similarly, the signal can be converted to the Audio Doppler format and sent to the speakers. Unlike CW Doppler, however, the transducer array does not have to be divided into two parts for continuous transmission and reception.

(Color Doppler)

- Spectral Doppler provides flow information along a fixed direction (CW) or within a single Doppler gate (PW). Color Doppler, on the other hand, provides real-time two-dimensional flow information. The ultrasonic data acquisition process in Color Doppler is similar to that in B-mode except that the ultrasonic pulse needs to be fired multiple times along the same direction sequentially before moving to the next beam. In other words, there are many Doppler gates (similar to the PW gate) along each ultrasonic beam and the flow profile within each gate is estimated. The detected flow information is then encoded in colors and superimposed on the two-dimensional B-mode image.



- In order to achieve sufficient frame rate for real-time imaging, the number of samples (i.e., the number of firings) used for flow estimation is typically limited to around 5 to 15. Compared to Spectral Doppler, where usually 32 to 128 point FFT is performed, the accuracy of flow estimation is limited and hence it is relatively a more qualitative tool for flow assessment. Spectral Doppler is often required for quantitative flow measurements.

- Similar to PW Doppler, the maximum flow velocity that can be detected without being aliased is limited to $\lambda/(4 \cdot PRI)$. Since there are multiple gates along an ultrasonic beam, the PRI effectively limits the maximum display depth. In other words, the maximum display depth cannot exceed $c \cdot PRI/2$, where c is the sound velocity.

- Instead of the previously described time-frequency analysis techniques (such as Fourier transform), more efficient flow estimation methods need to be employed in order to provide important aspects of blood flow with a limited number of samples. These important aspects include flow direction, mean velocity, flow turbulence and flow energy. In the following, an auto-correlation technique capable of estimating the above parameters will be described.

- Let $S(t)$ be the signal received from a particular gate, its auto-correlation function $R(t)$ is defined as

$$R(t) \equiv \int_{-\infty}^{\infty} S(t+\tau) S^*(\tau) d\tau$$

where * denotes the complex conjugate. In addition, the power spectrum $P(\omega)$ is the Fourier transform of $R(t)$ based on the Wiener-Khinchine's theorem, i.e.,

$$R(t) = \int_{-\infty}^{\infty} P(\omega) e^{j\omega t} d\omega$$

where $\omega = 2\pi f$ and f denotes the frequency. It then follows that the mean frequency (i.e., the first order moment) can be represented by

$$\overline{\omega} = \frac{\int_{-\infty}^{\infty} \omega P(\omega) d\omega}{\int_{-\infty}^{\infty} P(\omega) d\omega}.$$

Since

$$R(0) = \int_{-\infty}^{\infty} P(\omega) d\omega$$

and

$$-j \frac{dR(t)}{dt}\Big|_{t=0} = \int_{-\infty}^{\infty} \omega P(\omega) d\omega,$$

we have

$$\overline{\omega} = -j\frac{R'(0)}{R(0)}$$

where ' represents the time derivative. The above equation indicates that the mean Doppler shift frequency can be directly estimated using time domain signals without calculating the frequency spectrum.

- The above equation can be further simplified. Let $R(t) = |R(t)|e^{j\theta(t)}$, it is straightforward to see that $|R(t)|$ is an even function and $\theta(t)$ is an odd function. Using the above relations, the mean frequency $\overline{\omega}$ can be simplified as

$$\overline{\omega} = \theta'(0) \approx \frac{\theta(T) - \theta(0)}{T} = \frac{\theta(T)}{T}$$

where *T* is the time interval between two consecutive ultrasonic pulses (i.e., PRI). Note that $\theta(0) = 0$ since $\theta(t)$ is an odd function.

- In practice, $R(T)$ is simply the discrete auto-correlation function evaluated at the first lag. Since the number of firings per ultrasonic beam is limited, $R(T)$ is obtained by summing over a limited range. Suppose there are only two samples available (i.e., $S(0)$ and $S(T)$), $R(T) = S(T)S^*(0)$ and hence $\theta(T) = \angle S(T) - \angle S(0)$. In this case, the Doppler shift frequency is simply the phase change between two consecutive measurements divided by time T. Since the phase is related to the distance between the object and the transducer, the mean Doppler shift frequency also represents the change in distance during time *T*, which is indeed the definition of velocity.

- Flow direction is determined by the sign of mean frequency. A positive mean frequency means a flow is moving toward the transducer, a negative mean frequency means a flow is moving away from the transducer.

- The extent of turbulence in blood flow can be related to the variance of the power spectrum. In other words, the spectrum spread broadens in the presence of flow turbulence. The variance σ^2 is represented as the following:

$$\sigma^2 = \frac{\int_{-\infty}^{\infty} (\omega - \overline{\omega})^2 P(\omega)\, d\omega}{\int_{-\infty}^{\infty} P(\omega)\, d\omega} = \overline{\omega^2} - \overline{\omega}^2.$$

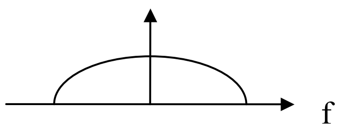
- The variance σ^2 can be further simplified as the following

$$\sigma^2 \approx \frac{2}{T^2}\left(1 - \frac{|R(T)|}{R(0)}\right).$$

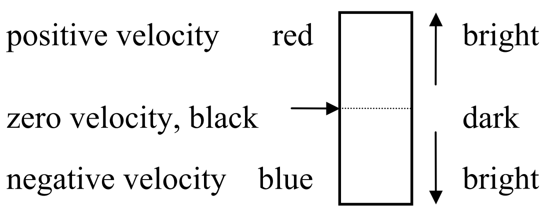
Therefore, the flow turbulence can also be estimated directly from the received time domain signals.

- The total energy E of the flow signal can be found by integrating the power spectrum over the entire frequency range, i.e., $E = \int_{-\infty}^{\infty} P(\omega)\, d\omega$. Since $\int_{-\infty}^{\infty} P(\omega)\, d\omega = R(0)$, the parameter E can also be estimated directly in time domain.

- Consider the following hypothetical flow, the mean frequency $\overline{\omega}$ becomes zero due to the symmetry. The energy E, however, is non-zero since it represents the total area under the spectrum. Therefore, the energy mode is sometimes preferred in some clinical situations, particularly when the direction of the flow is not of primary interest.

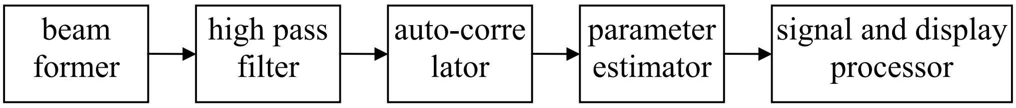


- As mentioned previously, these flow parameters are encoded in colors for the real-time display. The following example shows a typical way of mapping velocities (including directionality) to different colors.



- Two-dimensional maps can also be constructed to encode two different parameters into colors simultaneously. For example, one dimension may represent the velocity and the other may represent the variance.

- A simple block diagram for a Color Doppler processing chain is shown below. Note that the purpose of the high pass filter it the same as that in Spectral Doppler, i.e., it is a wall filter to filter out strong signals from non-moving and slow moving structures.



- Since multiple firings along an ultrasonic beam are required for flow estimation, the frame rate impact becomes a significant issue. In order to achieve sufficient frame rate to image flow dynamics, Color Doppler processing is usually confined to a small area of interest. Outside this area, only B-mode signals (i.e., single firing) are acquired. In addition, the beam spacing may be chosen higher than the Nyquist criterion to gain frame rate at the price of aliasing artifacts. In some high end systems, parallel beams are formed simultaneously in order to further increase frame rate.

- Doppler shift frequencies can also be used to estimate other types of motion, such as heart motion. In this case, motion abnormalities can be visualized by using color encoded, two-dimensional Doppler image in real time. Since the signal from heart muscle is sufficiently strong, the design of the wall filter can be greatly simplified.

- One inherent limitation of motion estimation using Doppler effect is that only motion parallel to the beam direction can be detected. Therefore, how to detect motion in two dimensions (or three dimensions, in general) is an important topic in diagnostic ultrasound.