

Discrete Transforms

A transform is simply another term for a mathematical mapping process. Most of the transforms in this lecture map the image data from the spatial domain to the spectral (frequency or sequence) domain, where all the pixels in the input spatial domain contribute to each value in the output spectral domain, as illustrated in Figure 71.

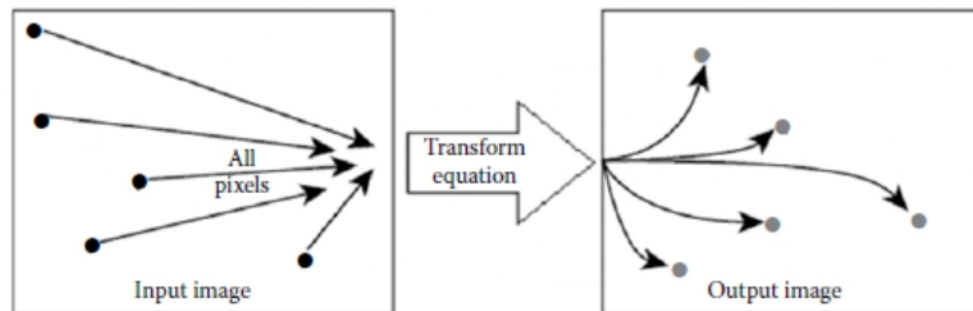


Figure 71. **Discrete transforms.** Spectral (frequency or sequence) transforms require all the values in the image to obtain each individual value in the spectrum. All pixels in the input image contribute to each value in the output image for spectral transforms.

These transforms are used as tools in many areas of engineering and science, including digital image processing. Originally defined in their continuous forms, they are commonly used today in their discrete (sampled) forms. The large number of arithmetic operations required for the discrete transforms, combined with the massive amounts of data in an image, requires a great deal of computer power. The ever increasing compute power, memory capacity, and fast disk storage available today make the use of these transforms much more feasible than in the past.

The discrete form of these transforms is created by sampling the continuous form of the functions on which these transforms are based, that is, the basis functions. The functions used for these transforms are typically sinusoidal or rectangular and the sampling process, for the one-dimensional (1D) case, provides us with basis vectors. When we extend these into two dimensions, as we do for images, they are basis matrices or basis images (see Figure 72). The process of transforming the image data into another domain, or mathematical space, amounts to projecting the image onto the basis images. The frequency transforms use the entire image, to discover spatial frequency information.

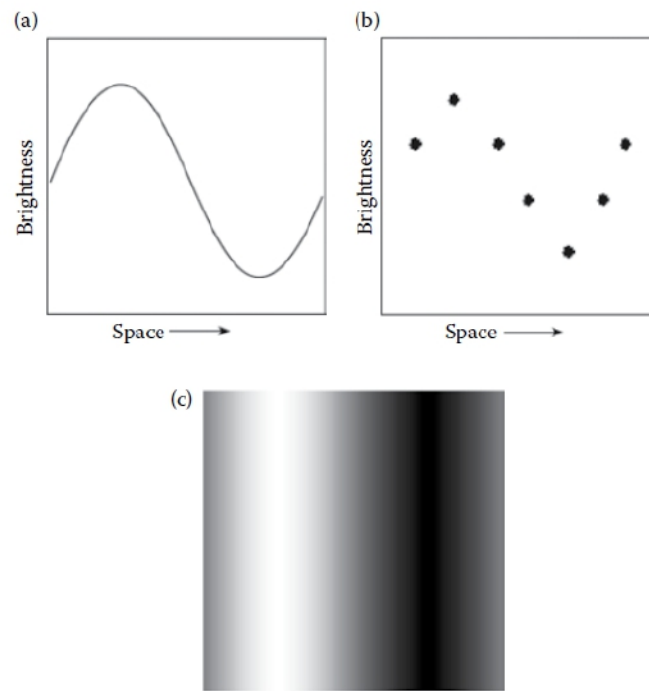


Figure 72. **Basis vectors and images.** (a) A basis function: a 1D sinusoid. (b) A basis vector: a sampled 1D sinusoid. (c) A basis image: A sampled sinusoid shown in 2D as an image. The pixel brightness in each row corresponds to the sampled values of the 1D sinusoid, which are repeated along each column. This is a horizontal sine wave of frequency 1.

The general form of the transformation equation, assuming an $M \times N$ image, is given by:

$$T(u, v) = k \sum_{r=0}^{M-1} \sum_{c=0}^{N-1} I(r, c) B(r, c; u, v)$$

Here, u and v are the frequency domain variables, k is a constant that is transform dependent, $T(u, v)$ are the transform coefficients, and $B(r, c; u, v)$ correspond to the basis images. The notation $B(r, c; u, v)$ defines a set of basis images, corresponding to each different value for u and v , and the size of each is $r \times c$ (Figure 73). The transform coefficients, $T(u, v)$, are the projections of $I(r, c)$ onto each $B(u, v)$, then scaled by the constant k . This is illustrated in Figure 74. These coefficients tell us how similar the image is to the basis image; the more alike they are, the bigger the coefficient. This transformation process amounts to decomposing the image into a weighted sum of the basis images, where the coefficients $T(u, v)$ are the weights.

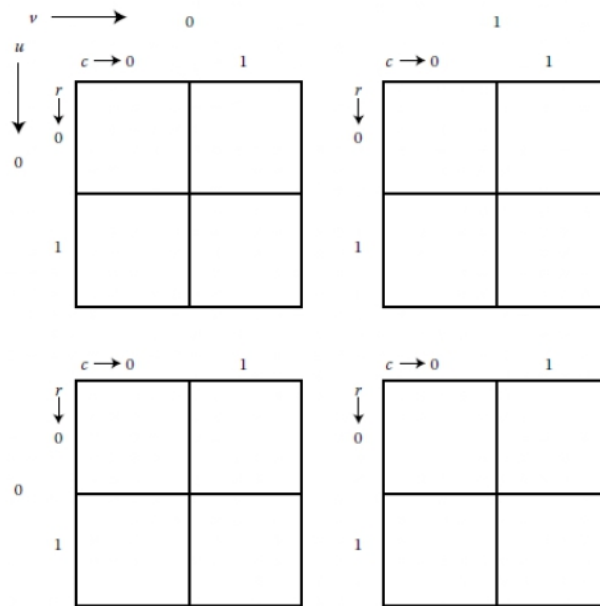


Figure 73. A set of basis images $B(r,c;u,v)$. Size and coordinates of generic basis images for a 2×2 transform

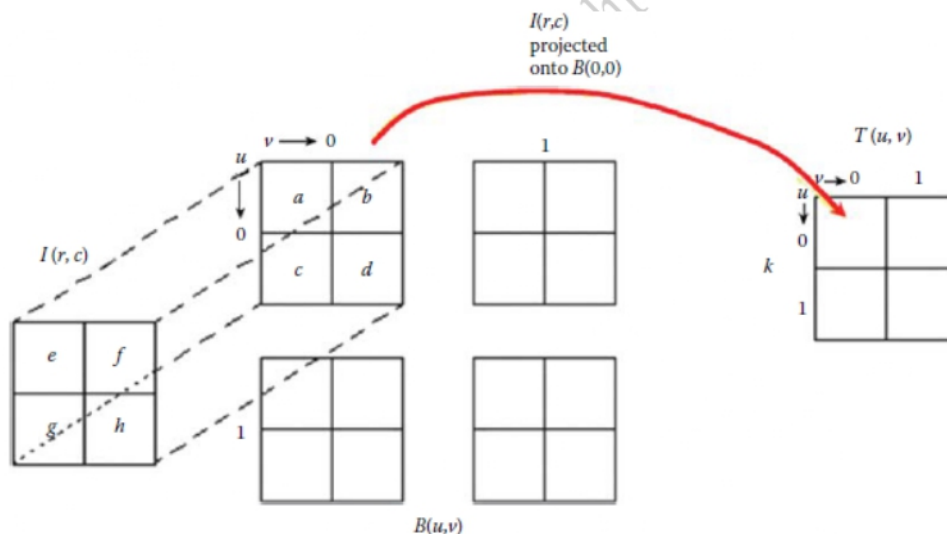


Figure 74. **Transform coefficients**. To find the transform coefficients, $T(u,v)$, we project the image, $I(r,c)$, onto the basis images, $B(u,v)$, and multiply by k . For example, $T(0,0)$ is the projection of $I(r,c)$ onto $B(0,0)$, which equals $(ea + fb + gc + hd)$.

Example 1.

Let

$$I(r,c) = \begin{bmatrix} 5 & 3 \\ 1 & 2 \end{bmatrix}$$

And, let

$$B(u,v,r,c) = \begin{bmatrix} \begin{bmatrix} +1 & +1 \\ +1 & +1 \end{bmatrix} & \begin{bmatrix} +1 & -1 \\ +1 & -1 \end{bmatrix} \\ \begin{bmatrix} +1 & +1 \\ -1 & -1 \end{bmatrix} & \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix} \end{bmatrix}$$

Then,

$$T(u,v) = k \begin{bmatrix} T(0,0) = 5(1) + 3(1) + 1(1) + 2(1) = 11 \\ T(0,1) = 5(1) + 3(-1) + 1(1) + 2(-1) = 1 \\ T(1,0) = 5(1) + 3(1) + 1(-1) + 2(-1) = 5 \\ T(1,1) = 5(1) + 3(-1) + 1(-1) + 2(1) = 3 \end{bmatrix} = k \begin{bmatrix} 11 & 1 \\ 5 & 3 \end{bmatrix}$$

To obtain the image from the transform coefficients, we apply the inverse transform equation:

$$I(r,c) = T^{-1}[T(u,v)] = k' \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u,v) B^{-1}(r,c;u,v)$$

Here, the $T^{-1}[T(u,v)]$ represents the inverse transform, and the $B^{-1}(r,c;u,v)$ represents the inverse basis images, and k' is a constant that is transform dependent.

Example 2

From the previous example, we have:

$$T(u,v) = k \begin{bmatrix} 11 & 1 \\ 5 & 3 \end{bmatrix}$$

And, let

$$B^{-1}(u,v,r,c) = \begin{bmatrix} \begin{bmatrix} +1 & +1 \\ +1 & +1 \end{bmatrix} & \begin{bmatrix} +1 & -1 \\ +1 & -1 \end{bmatrix} \\ \begin{bmatrix} +1 & +1 \\ -1 & -1 \end{bmatrix} & \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix} \end{bmatrix}$$

Then,

$$I(r,c) = \begin{bmatrix} I(0,0) = 11(1) + 1(1) + 5(1) + 3(1) = 20 \\ I(0,1) = 11(1) + 1(-1) + 5(1) + 3(-1) = 12 \\ I(1,0) = 11(1) + 1(1) + 5(-1) + 3(-1) = 4 \\ I(1,1) = 11(1) + 1(-1) + 5(-1) + 3(1) = 8 \end{bmatrix} = k' \begin{bmatrix} 20 & 12 \\ 4 & 8 \end{bmatrix}$$

Is this correct? It depends on k and k' , since:

$$I(r,c) = \begin{bmatrix} 5 & 3 \\ 1 & 2 \end{bmatrix}$$

Comparing our results we see that we must multiply our answer by $1/4$. What does this tell us?—It tells us that to complete the transform pair, $B(u,v;r,c)$ and $B^{-1}(u,v;r,c)$, we need to define k and k' . We need to be able to recover our original image to have a proper transform pair. We can solve this by letting $k' = 1/4$, or by letting $k = k' = 1/2$. Note that $1/2$ will normalize the magnitude of the basis images to 1. Remember that the magnitude of a vector can be found by the square root of the sum of the squares of the vector components; in this case, magnitude of the basis images $= \sqrt{1^2 + (\pm 1)^2 + (\pm 1)^2 + (\pm 1)^2} = \sqrt{4} = 2$. Therefore, to normalize the magnitude to 1, we need to divide by 2, or multiply by $1/2$.

Fourier Transform

The Fourier transform is the best known, and the most widely used, of the transforms. It was developed by Jean Baptiste Joseph Fourier (1768–1830) to explain the distribution of temperature and heat conduction. The Fourier transform has found numerous uses, including vibration analysis in mechanical engineering, circuit analysis in electrical engineering, and in digital image processing. The Fourier transform decomposes a complex signal into a weighted sum of a zero frequency term (the DC term that is related to the average value), and sinusoidal terms, the basis functions, where each sinusoid is a harmonic of the fundamental. Fourier transform theory begins with the 1D continuous transform, defined as follows:

$$F(v) = \int_{-\infty}^{\infty} I(c)e^{-j2\pi vc} dc$$

The basis functions, $e^{-j2\pi vc}$, are complex exponentials, and sinusoidal in nature. Also, the continuous Fourier transform theory assumes that the functions start at $-\infty$ and go to $+\infty$, so they are continuous and everywhere. This aspect of the underlying theory is important for the periodic property of the Fourier transform

Example 3

Given the simple rectangle function shown in **Figure 75a**, we can find the Fourier transform by applying the equation defined above.

$$\begin{aligned} F(v) &= \int_{-\infty}^{\infty} I(c)e^{-j2\pi vc} dc \\ &= \int_0^C Ae^{-j2\pi vc} dc \\ &= -\frac{A}{j2\pi v} [e^{-j2\pi vC}]_0^C = \frac{-A}{j2\pi v} [e^{-j2\pi vC} - 1] \\ &= \frac{A}{j2\pi v} [e^{j\pi vC} - e^{-j\pi vC}] e^{-j\pi vC} \end{aligned}$$

then we use the trigonometric identity, $\sin \theta = (e^{j\theta} - e^{-j\theta})/2j$

$$= \frac{A}{\pi v} \sin(\pi vC) e^{-j\pi vC}$$

This result is a complex function, and here we are interested in the magnitude (defined in the next section), which is

$$|F(v)| = \left| \frac{A}{\pi v} \right| |\sin(\pi vC)| |e^{-j\pi vC}|$$

Now, we multiply through by C/C , and the magnitude of $e^{-j\pi vC} = 1$, we can get it in the form of a sinc function:

$$= AC \left| \frac{\sin(\pi vC)}{(\pi vC)} \right| = AC |\text{sinc}(\nu C)|$$

Figure 75b shows this result.

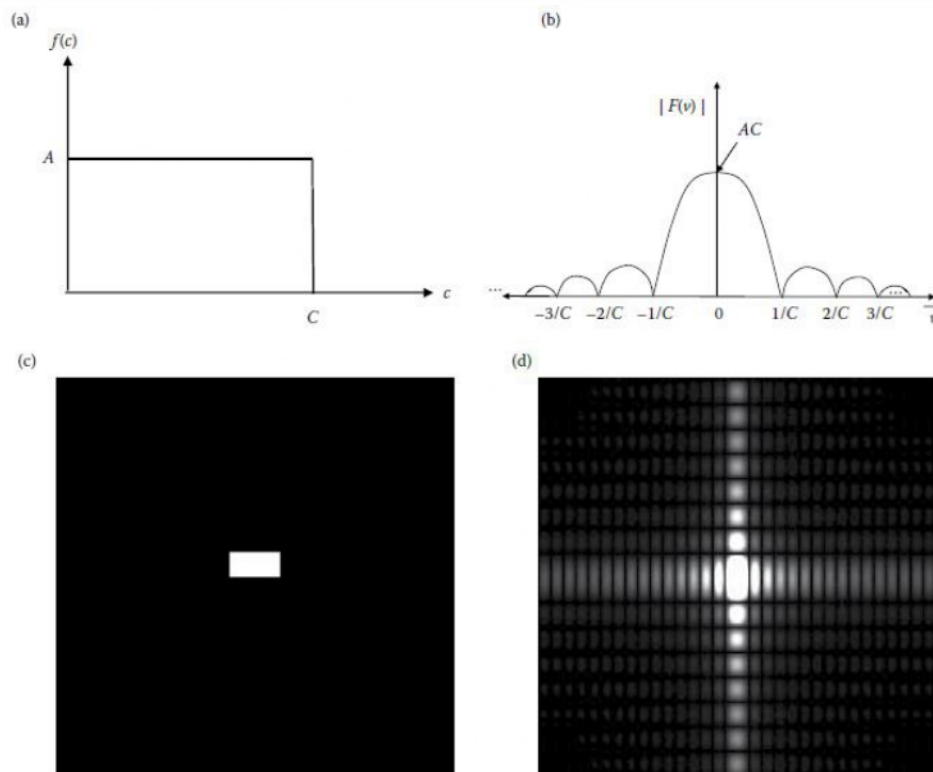


Figure 75. **Fourier transform example.** (a) The one-dimensional rectangle function, (b) the magnitude of Fourier transform of the 1D rectangle function: $|F(v)| = AC \left| \frac{\sin(\pi vC)}{\pi vC} \right| = AC |\text{sinc}(vC)|$. (c) Two-dimensional rectangle function as an image, (d) the magnitude of the Fourier transform, called the Fourier spectrum, of the 2D rectangle.

Figure 75c shows the 2D rectangle function, with the brightness of the image representing the magnitude of the function. In Figure 75d, the magnitude of the Fourier spectrum in image form. It is customary to display the magnitude only of a Fourier spectrum, as the Fourier transform contains complex terms, which have real and imaginary parts. The magnitude is however a real quantity and is also referred to as the Fourier spectrum or frequency spectrum.

The reasons for introducing this example are as follows: (1) to illustrate the continuous and infinite nature of the basis functions in the underlying theory, (2) to illustrate that when we have a function that ends abruptly in one domain, such as the function $F(c)$, it leads to a continuous series of decaying ripples in the other domain as shown in Figure 75b and d, and (3) to show that the width of the rectangle in one domain is inversely proportional to the spacing of the ripples in the other domain. This will be useful in understanding the nature of phenomena that occurs in images at object boundaries, especially when we apply filters; but first, we will explain the details of the discrete Fourier transform (DFT).

The One-Dimensional Discrete Fourier Transform

The equation for the 1D DFT is:

$$F(v) = \frac{1}{N} \sum_{c=0}^{N-1} I(c) e^{-j2\pi vc/N}$$

The inverse DFT is given by:

$$F^{-1}[F(v)] = I(c) = \sum_{v=0}^{N-1} F(v) e^{j2\pi vc/N}$$

Where the $F^{-1}[\]$ notation represents the inverse transform. These equations correspond to one row of an image; note that as we move across a row, the column coordinate is the one that changes. The base of the natural logarithmic function, e , is about 2.71828; j , the imaginary coordinate for a complex number, equals $\sqrt{-1}$. The basis functions are sinusoidal in nature, as can be seen by Euler's identity:

$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$

Putting this equation into the DFT equation by substituting $\theta = -2\pi vc/N$, and remembering that $\cos(\theta) = \cos(-\theta)$ and $\sin(-\theta) = -\sin(\theta)$, the 1D DFT equation can be written as:

$$F(v) = \frac{1}{N} \sum_{c=0}^{N-1} I(c) [\cos(2\pi vc/N) - j\sin(2\pi vc/N)] = \text{Re}(v) + j\text{Im}(v)$$

The $F(v)$ is also complex, with the real part corresponding to the cosine terms, and the imaginary part corresponding to the sine terms. If we represent a complex spectral component by $F(v) = \text{Re}(v) + j\text{Im}(v)$, where $\text{Re}(v)$ is the real part and $\text{Im}(v)$ is the imaginary part, then we can define the magnitude and phase of a complex spectral component as:

$$\text{Magnitude} = |F(v)| = \sqrt{[\text{Re}(v)]^2 + [\text{Im}(v)]^2}$$

and

$$\text{Phase} = \phi(v) = \tan^{-1} \left[\frac{\text{Im}(v)}{\text{Re}(v)} \right]$$

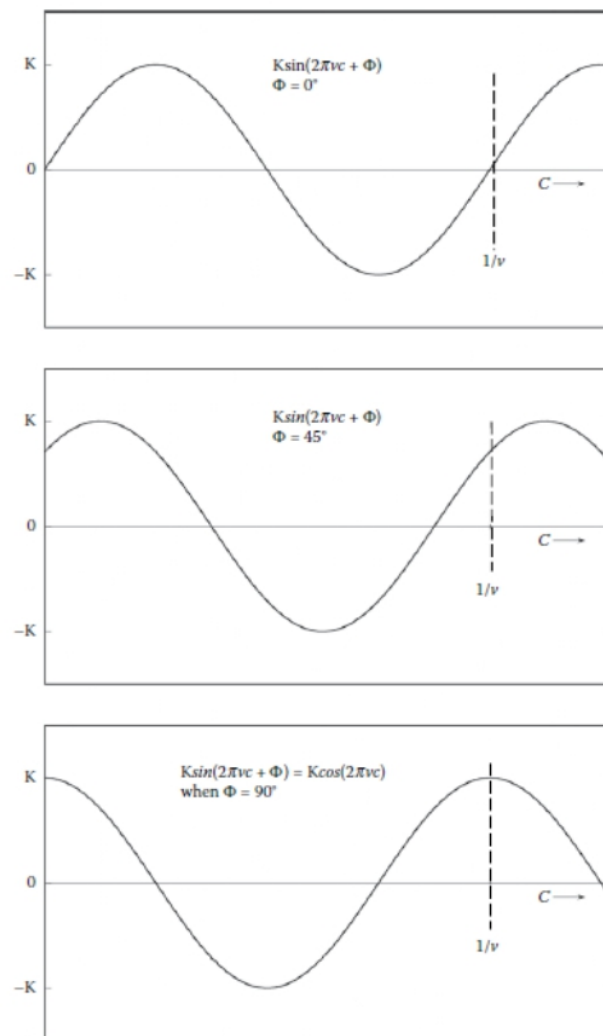
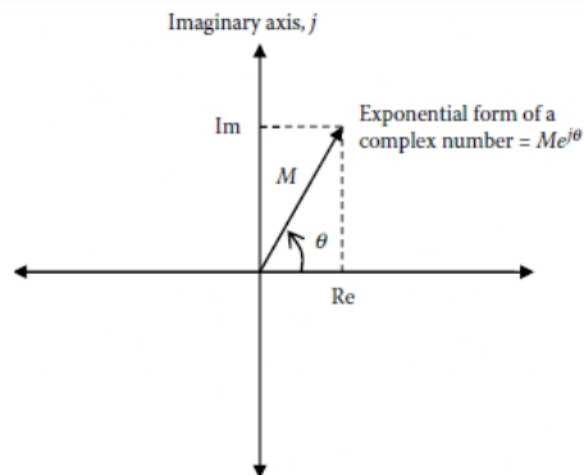
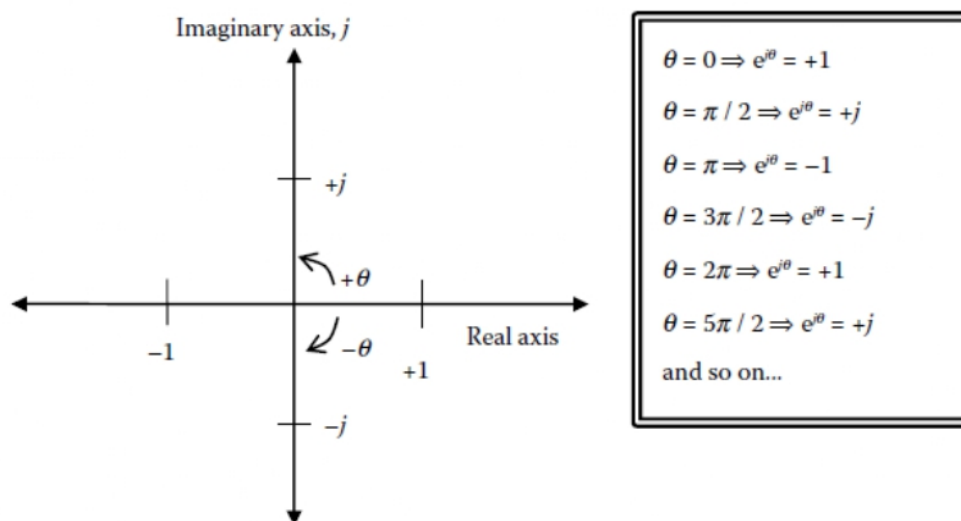


Figure 76. Magnitude and phase of sinusoidal waves.

The magnitude of a sinusoid is simply its peak value, and the phase determines where the origin is, or where the sinusoid starts (see Figure 76). The basis functions are simply sinusoids at varying frequencies, the complex exponential notation, e^{jx} , is simply a mathematical notational tool to make it easier to write and manipulate the equations. In Figure 77, we see that a complex number can be expressed in rectangular form, described by the real and imaginary parts; or in exponential form, by the magnitude and phase. A memory aid for evaluating $e^{j\theta}$ is given in Figure 78.

FIGURE 77. **Complex numbers**Figure 78. **A memory aid for evaluating $e^{j\theta}$.** The angle is measured from the real axis counterclockwise.**Example 4**

Given $I(c) = [3, 2, 2, 1]$, corresponding to the brightness values of one row of a digital image. Find $F(v)$ in both rectangular and exponential forms.

$$F(v) = \frac{1}{N} \sum_{c=0}^{N-1} I(c) e^{-j2\pi vc/N}$$

$$F(0) = \frac{1}{4} \sum_{c=0}^3 I(c) e^{-j2\pi(0)c/4} = \frac{1}{4} \sum_{c=0}^3 I(c) e^0 = \frac{1}{4} [I(0) + I(1) + I(2) + I(3)] = \frac{1}{4} [3 + 2 + 2 + 1] = 2$$

$$F(1) = \frac{1}{4} \sum_{c=0}^3 I(c) e^{-j2\pi(1)c/4} = \frac{1}{4} [3e^0 + 2e^{-j\pi/2} + 2e^{-j\pi} + 1e^{-j3\pi/2}] = \frac{1}{4} [3 + 2(-j) + 2(-1) + 1(j)] = \frac{1}{4} [1 - j]$$

$$F(2) = \frac{1}{4} \sum_{c=0}^3 I(c) e^{-j2\pi(2)c/4} = \frac{1}{4} [3e^0 + 2e^{-j\pi} + 2e^{-j2\pi} + 1e^{-j3\pi}] = \frac{1}{4} [3 + (-2) + 2 + (-1)] = \frac{1}{2}$$

$$F(3) = \frac{1}{4} \sum_{c=0}^3 I(c) e^{-j2\pi(3)c/4} = \frac{1}{4} [3e^0 + 2e^{-j3\pi/2} + 2e^{-j3\pi} + 1e^{-j9\pi/2}] = \frac{1}{4} [3 + 2j + 2(-1) + 1(-j)] = \frac{1}{4} [1 + j]$$

Therefore, we have:

$$F(v) = \left[2, \frac{1}{4}[1 - j], \frac{1}{2}, \frac{1}{4}[1 + j] \right]$$

Next, put these into the exponential form.

$$F(0) = 2 = 2 + 0j \Rightarrow M = \sqrt{2^2 + 0^2} = 2; \theta = \tan^{-1} \left[\frac{0}{2} \right] = 0$$

$$F(1) = \frac{1}{4}[1 - j] = \frac{1}{4} - \frac{1}{4}j \Rightarrow M = \sqrt{\left(\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2} \cong 0.35; \theta = \tan^{-1} \left[\frac{-\frac{1}{4}}{\frac{1}{4}} \right] = -\pi/4$$

$$F(2) = 1/2 = 1/2 + 0j \Rightarrow M = \sqrt{(1/2)^2 + 0^2} = 0.5; \theta = \tan^{-1} \left[\frac{0}{1/2} \right] = 0$$

$$F(3) = \frac{1}{4}[1 + j] = \frac{1}{4} + \frac{1}{4}j \Rightarrow M = \sqrt{\left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2} \cong 0.35; \theta = \tan^{-1} \left[\frac{\frac{1}{4}}{\frac{1}{4}} \right] = \pi/4$$

Therefore, we have:

$$F(v) = [2, 0.35e^{-j\pi/4}, 0.5, 0.35e^{j\pi/4}]$$

The Two-Dimensional Discrete Fourier Transform

Extending the DFT to the 2D case for images, we can decompose an image into a weighted sum of 2D sinusoidal terms. The physical interpretation of a 2D sinusoid is shown in Figure 79. The sinusoid that is not directly on the u - or v -axis can be broken down into separate frequency terms by finding the period along each axis. Assuming a square $N \times N$ image, the equation for the 2D DFT is:

$$F(u, v) = \frac{1}{N} \sum_{r=0}^{N-1} \sum_{c=0}^{N-1} I(r, c) e^{-j2\pi(ur+vc)/N}$$

The Fourier transform equation as:

$$F(u, v) = \frac{1}{N} \sum_{r=0}^{N-1} \sum_{c=0}^{N-1} I(r, c) \left[\cos\left(\frac{2\pi}{N}(ur + vc)\right) - j \sin\left(\frac{2\pi}{N}(ur + vc)\right) \right]$$

The $F(u, v)$ is also complex, with the real part corresponding to the cosine terms, and the imaginary part corresponding to the sine terms. If we represent a complex spectral component by $F(u, v) = \text{Re}(u, v) + j\text{Im}(u, v)$, where $\text{Re}(u, v)$ is the real part and $\text{Im}(u, v)$ is the imaginary part, then the magnitude and phase of a complex spectral component can be defined as:

$$\text{Magnitude} = |F(u, v)| = \sqrt{[\text{Re}(u, v)]^2 + [\text{Im}(u, v)]^2}$$

and

$$\text{Phase} = \phi(u, v) = \tan^{-1} \left[\frac{\text{Im}(u, v)}{\text{Re}(u, v)} \right]$$

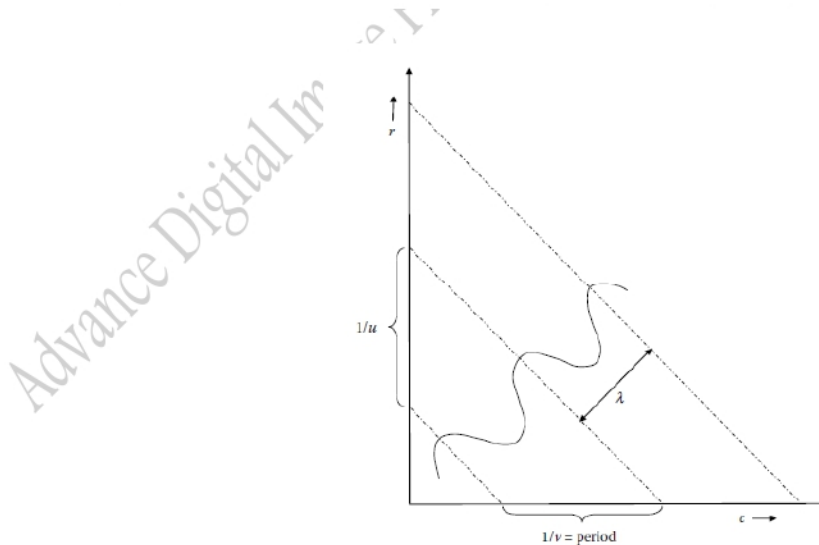


Figure 79. Physical interpretation of a two-dimensional sinusoid. The wavelength of the sinusoid is: $\lambda = 1/2\pi uv$, where (u, v) are the frequencies along (r, c) , and the periods are $1/u$ and $1/v$.

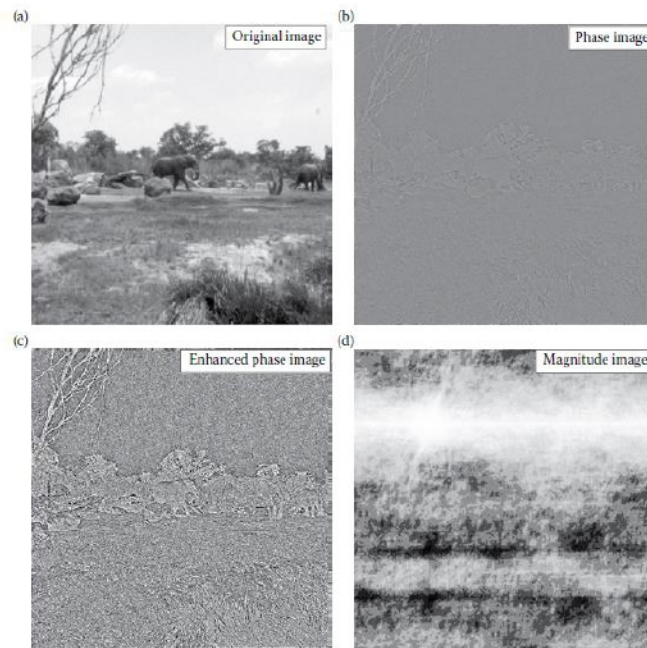


Figure 80. **Fourier transform phase and magnitude image information.** (a) Original image, (b) phase only image, (c) contrast enhanced version of image (b) to show detail, (d) magnitude only image after histogram equalization. The phase only image is created by taking a Fourier transform, setting all the magnitudes equal to 1, and performing an inverse Fourier transform. The magnitude only image is created by taking a Fourier transform, setting the phase to a fixed value, such as 0, then performing an inverse Fourier transform

Figure 80 shows images recovered with the phase or magnitude only. With phase only, we lose the relative magnitudes, which results in a loss of contrast (see Figure 80b), but we retain the relative placement of objects—in other words, the phase data contains information about where objects are in an image. With the magnitude only image, we retain the contrast, but lose all the important detail, which is essential to image understanding. After the transform is performed, to get our original image back, we need to apply the inverse transform. The inverse 2D DFT is given by:

$$F^{-1}[F(u,v)] = I(r,c) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) e^{j2\pi(ur+vc)/N}$$

The $F^{-1}[\]$ notation represents the inverse transform.

One important property of the Fourier transform is called **separability**, which means that the 2D basis image can be decomposed into two product terms, where each term depends only on the rows or columns. Also, if the basis images are separable, then the 2D result can be found by successive application of the 1D transform. This is illustrated by first separating the basis image term (also called the transform kernel) into a product, as follows.

$$e^{-j2\pi(ur+vc)/N} = e^{-j2\pi ur/N} e^{-j2\pi vc/N}$$

Fourier transform equation in the following form.

$$F(u, v) = \frac{1}{N} \sum_{r=0}^{N-1} \left(e^{-j2\pi ur/N} \right) \sum_{c=0}^{N-1} I(r, c) e^{-j2\pi vc/N}$$

Fourier Transform Properties

1. Linearity

The Fourier transform is a linear operator and is shown by the following equations.

$$\begin{aligned} F[aI_1(r, c) + bI_2(r, c)] &= aF_1(u, v) + bF_2(u, v) \\ aI_1(r, c) + bI_2(r, c) &= F^{-1}[aF_1(u, v) + bF_2(u, v)] \end{aligned}$$

where a and b are constants.

2. Convolution

Convolution in one domain is the equivalent of multiplication in the other domain; this is what allows us to perform filtering in the spatial domain with convolution masks. Using * to denote the convolution operation, and F[] for the forward Fourier transform, and F⁻¹[] for the inverse Fourier transform, these equations define this property:

$$\begin{aligned} F[I_1(r, c) * I_2(r, c)] &= F_1(u, v) F_2(u, v) \\ I_1(r, c) * I_2(r, c) &= F^{-1}[F_1(u, v) F_2(u, v)] \\ F[I_1(r, c) I_2(r, c)] &= F_1(u, v) * F_2(u, v) \\ I_1(r, c) I_2(r, c) &= F^{-1}[F_1(u, v) * F_2(u, v)] \end{aligned}$$

3. Translation

The translation property of the Fourier transform is given by the following equations:

$$F[I(r - r_0, c - c_0)] = F(u, v)e^{-j2\pi(ur_0 + vc_0)/N}$$

$$I(r - r_0, c - c_0) = F^{-1}[F(u, v)e^{-j2\pi(ur_0 + vc_0)/N}]$$

These equations tell us that if the image is moved, the resulting Fourier spectrum undergoes a phase shift, but the magnitude of the spectrum remains the same.

4. Modulation

The modulation property, also called the frequency translation property, is given by:

$$F[I(r, c)e^{j2\pi(u_0r + v_0c)/N}] = F(u - u_0, v - v_0)$$

$$I(r, c)e^{j2\pi(u_0r + v_0c)/N} = F^{-1}[F(u - u_0, v - v_0)]$$

If the image is multiplied by a complex exponential (remember this is really a form of a sinusoid), its corresponding spectrum is shifted. This property is illustrated in Figure 81.

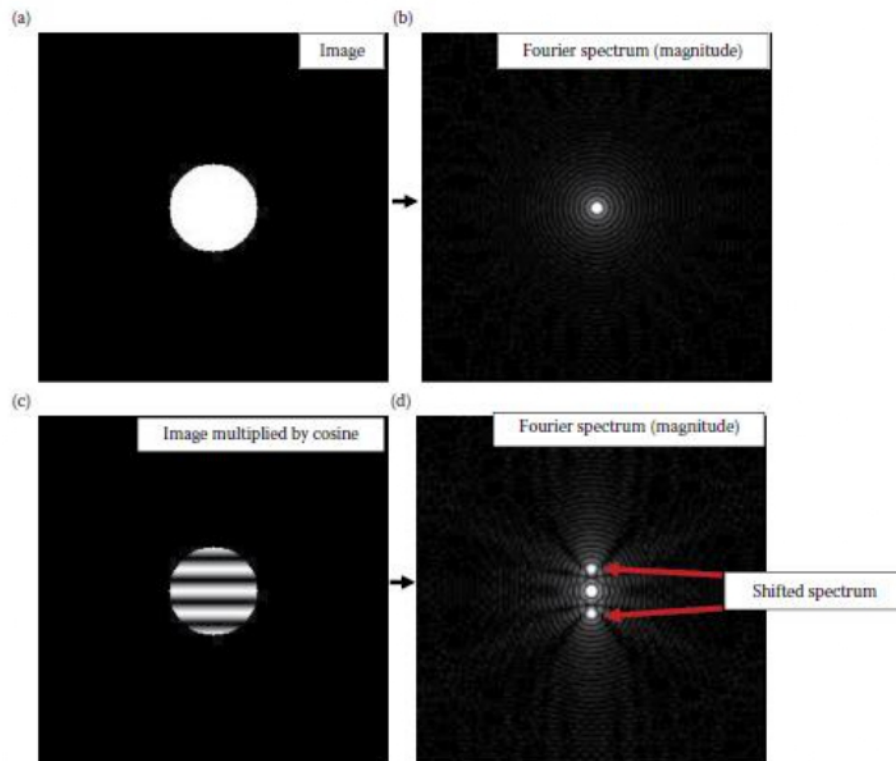


Figure 81. Modulation property results in frequency shift. (a) Original image, (b) magnitude of Fourier spectrum of (a) represented as an image, (c) original image multiplied by a vertical cosine wave at a relative frequency of 16 (16 cycles per image), (d) magnitude of Fourier spectrum of (c) represented as an image. Note that the spectrum has been shifted by 16 above and below the origin (in these spectral images the origin is in the center of the image).

5. Rotation

The rotation property is best illustrated by using polar coordinates.

$$r = x \cos(\theta), c = x \sin(\theta)$$

$$u = w \cos(\phi), v = w \sin(\phi)$$

The Fourier transform pair $I(r,c)$ and $F(u,v)$ become $I(x, \theta)$ and $F(w, \phi)$, respectively, and we can write a Fourier transform pair to illustrate the rotation property as follows:

$$I(x, \theta + \theta_0) = F^{-1} [F(w, \phi + \theta_0)]$$

$$F[I(x, \theta + \theta_0)] = F(w, \phi + \theta_0)$$

If the image is rotated by an angle θ^0 , then $F(u,v)$ is rotated by the same angle, and vice versa.

6. Periodicity

The DFT is periodic with period N , for an $N \times N$ image. This means:

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N) \dots$$

This is shown in Figure 82a. This figure shows nine periods, but the theoretical implication is that it continues in all directions to infinity. This property defines the implied symmetry in the discrete Fourier spectrum, and allows us to easily shift the spectrum so that the DC term, $F(0,0)$, originally in the upper left corner, is in the center of the image.

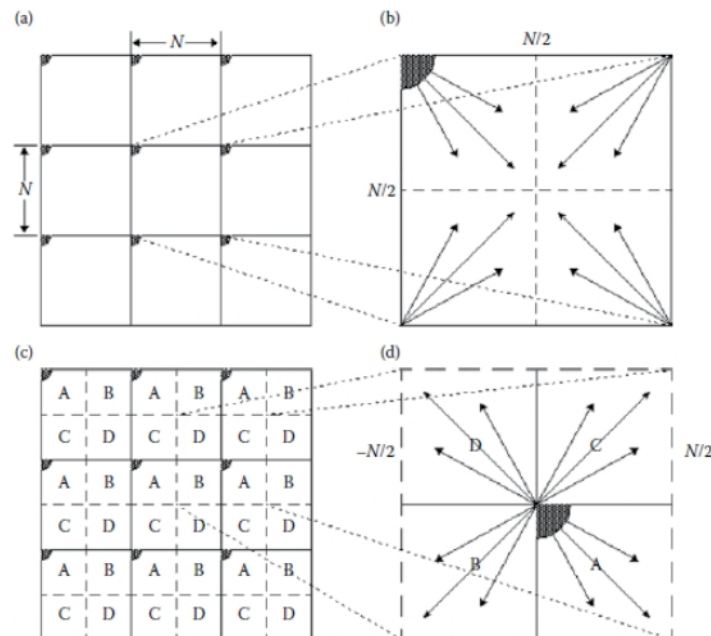


Figure 82. **Periodicity and discrete Fourier transform symmetry.** The periodicity property of the discrete Fourier spectrum allows us to easily shift the spectrum so that the DC term, $F(0,0)$, appears in the center. The shaded area represents the origin at $(0,0)$ (note: at $(1,1)$ in Matlab). (a) Implied symmetry with origin in upper-left corner. Each $N \times N$ block represents all the transform coefficients, and is repeated infinitely in all directions. (b) Increasing frequency in direction of arrows. (c) Periodic spectrum, with quadrants labeled A, B, C, D. (d) Spectrum shifted to center. Frequency increases in all directions as we move away from the origin.

7. Sampling and Aliasing

Sampling theory in digital signal processing states that we must sample a continuous signal with a sampling rate that is at least twice the highest frequency contained in the signal. This sampling rate is called the Nyquist rate. If the Nyquist rate is violated, aliasing occurs. Aliasing occurs when the continuous signal is sampled at a rate less than twice the highest frequency in the signal and appears as false information in the high frequencies.

Figure 83 illustrates the concept of aliasing. For simplicity, we will look at two periods and only consider the frequencies in the horizontal direction. By applying the periodicity of the Fourier transform, which says it is periodic with a period of $N \times N$ for an $N \times N$ image, if there is information in the signal at frequencies greater than $N/2$, overlap will occur. This overlap causes the aliasing, which manifests itself as false information in the high frequencies. In Figure 83a, we show the spectrum of two periods where the Nyquist criteria have been met. Figure 83b shows what happens if the original signal actually contains frequencies greater than $N/2$ in the horizontal direction: adjacent periods overlap and aliasing occurs. Note that one method to avoid aliasing is to band limit the original signal with a low-pass filter so that it does not contain any frequencies above the Nyquist frequency.

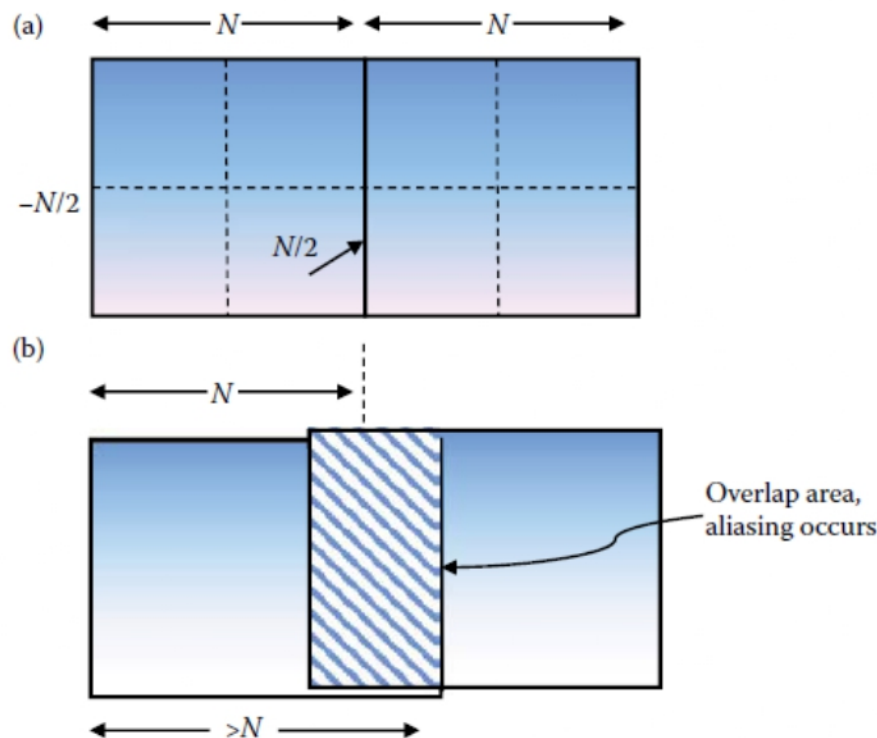


Figure 83. **Spectral aliasing.** (a) In two periods of the Fourier spectrum of an $N \times N$ image, sampled by the Nyquist rate, no frequencies in the signal are greater than $N/2$. (b) Two periods of the Fourier spectrum of an $N \times N$ image, sampled by a rate less than the Nyquist rate. The period implied is still $N \times N$, from $-N/2$ to $N/2$, but there are actual frequencies in the original image greater than $N/2$. In this case, the periods overlap causing false frequency information in the high frequencies. Since one period overlaps the next period, we get contributions from both, which creates false information.

Displaying the Discrete Fourier Spectrum

a spectral image is actually the magnitude data that has been remapped in a way that makes visualization easier. For displaying the magnitude of the Fourier spectrum, we usually shift the origin to the center. Applying the periodicity property and the modulation property with $u_0 = v_0 = N/2$, we obtain:

$$\begin{aligned}
 I(r, c) e^{j2\pi(Nr/2 + Nc/2)/N} \\
 &= I(r, c) e^{j\pi(r+c)} \\
 &= I(r, c) (-1)^{(r+c)}
 \end{aligned}$$

We can shift the spectrum by $N/2$ by multiplying the original image by $(-1)^{(r+c)}$, which will shift the origin to the center of the image (shown in Figure 82), it is easier to

understand the spectral information with the origin in the center and frequency increasing from the center out toward the edges.

The actual dynamic range of the Fourier spectrum is much greater than the 256 brightness levels (8-bits) available per color band with most image display devices. When we remap it to 256 levels, we can only see the largest values. We can apply contrast enhancement techniques to show more information, but we are still missing much of the visual information due to the limited dynamic range. To take advantage of the human visual system's response to brightness, we can greatly enhance the visual information available by displaying the following log transform of the spectrum:

$$\log(u, v) = k \log_{10}[1 + |F(u, v)|]$$

The log function compresses the data, and the scaling factor k remaps the data to the 0–255 range. Figure 84 shows the comparison of displaying the magnitude of the spectrum by direct remapping and contrast enhancement versus the log remap method. The log remap method shows much more information visually.

We learned that a function that ends abruptly in one domain results in rippling in the other domain that corresponds to the edges. Figure 85 displays images of simple geometric shapes and their corresponding spectral images. In addition to the magnitude information, the phase information is available in the Fourier spectrum. This information is not displayed as an image, but we have found it useful to illustrate phase changes as illustrated in translation property results in a phase shift of the spectrum. The phase information has a range of 0–360°, or 0– 2π radians.

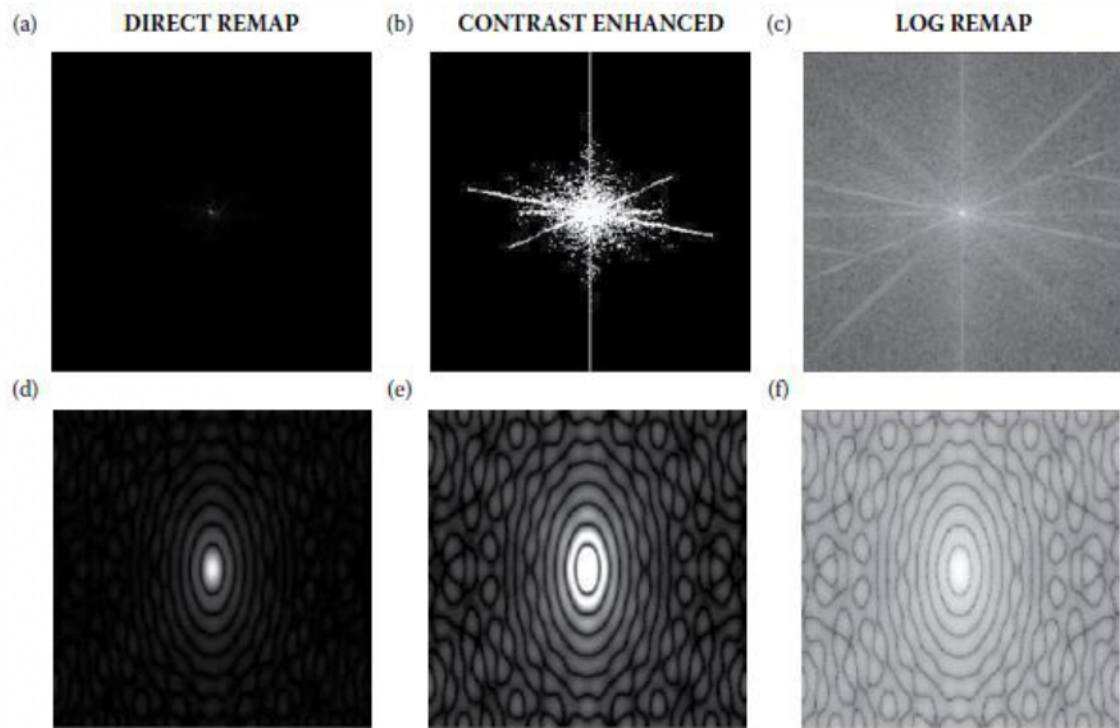


Figure 84. **Displaying DFT spectrum with various remap methods.** (a) Fourier magnitude spectrum of cam.pgm, direct remap to byte, (b) contrast-enhanced version of (a), (c) log remapped version of cam.pgm DFT spectrum, (d) Fourier magnitude spectrum of an ellipse, direct remap to byte, (e) contrast enhanced version of (d), (f) log remapped version of an ellipse DFT spectrum,

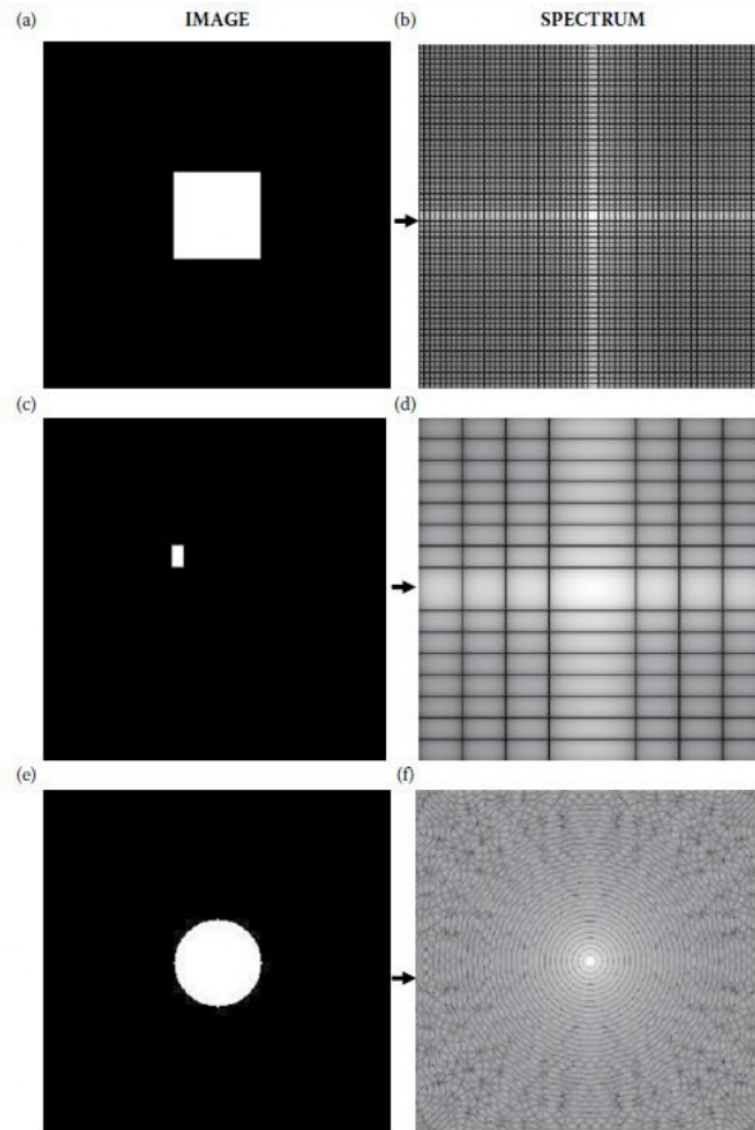


Figure 85. **Images of simple geometric shapes and their Fourier spectral images.** (a) An image of square, (b) the log remapped spectrum of the square, (c) a small rectangle, (d) the log remapped spectrum of the small rectangle, (e) an image of a circle, (f) the log remapped spectrum of the circle image

Discrete Cosine Transform

The cosine transform, like the Fourier transform, uses sinusoidal basis functions. The difference is that the cosine transform basis functions are not complex; they use only cosine and not sine functions. Since this transform uses only the cosine function, it can be calculated using only real arithmetic, instead of complex arithmetic as the DFT requires. The cosine transform is often used in image compression, as the Joint Photographers Expert Group (JPEG) image compression method.