2010 Spring (summer intensive) semester Special Lecture of Informatics

Topic 1. Image compression by orthogonal transformation

Session 3. (3) Discrete Fourier transformation and discrete cosine transformation

The framework of orthogonal and unitary transformations operating separately on rows and columns of an image matrix was explained in the previous session. An example of the orthogonal and unitary transformations as well as image compression by the transformations are presented today, in the final session of this topic.

Two-dimensional discrete Fourier transformation

In Session 2 of this course, Fourier transformation of a two-dimensional function was defined as

$$F(\nu_x, \nu_y) = \iint_{-\infty}^{\infty} f(x, y) \exp\{-i2\pi(\nu_x x + \nu_y y\} dxdy. \tag{1}$$

Since this equation can be rewritten to the following:

$$\begin{aligned} F(\nu_x, \nu_y) &= \iint_{-\infty}^{\infty} f(x, y) \exp(-i2\pi\nu_x x) \exp(-i2\pi\nu_y y) dxdy \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x, y) \exp(-i2\pi\nu_x x) dx \right] \exp(-i2\pi\nu_y y) dy, \end{aligned} \tag{2}$$

it corresponds to two one-dimensional Fourier transformations along the x-direction and y-direction, respectively. In Session 3 of this course, it was explained that one-dimensional discrete Fourier transformation of $u(n)$, an N-term sequence, was defined as follows:

$$U(k) = \sum_{n=0}^{N-1} u(n) \exp(-i2\pi \frac{k}{N} n) \quad (k = 0, 1, \ldots, N-1). \tag{3}$$

It follows from the above equation that the two-dimensional discrete Fourier transformation of $u(m, n)$, which is an array containing M terms of variable m and N terms of variable n, is defined as follows:

$$U(k, l) = \sum_{n=0}^{N-1} \left[\sum_{m=0}^{M-1} u(m, n) \exp(-i2\pi \frac{k}{M} m) \right] \exp(-i2\pi \frac{l}{N} n) \tag{4}$$

$(k = 0, 1, \ldots, M-1, l = 0, 1, \ldots, N-1)$. In case $M = N$, Eq. (4) is rewritten to

$$U(k, l) = \sum_{n=0}^{N-1} \left[\sum_{m=0}^{N-1} u(m, n) \exp(-i2\pi \frac{k}{N} m) \right] \exp(-i2\pi \frac{l}{N} n) \quad (k, l = 0, 1, \ldots, N-1), \tag{5}$$

i. e. Fourier transformation of a square digital image.

It was explained in the previous session that a separable unitary transformation of a matrix X can be expressed using a unitary matrix R as follows:

$$Z = RXR' \tag{6}$$

Now we express the transformation of Eq. (5) in the form of Eq. (6). Comparing the rows and columns in Eq. (6) and the variables in Eq. (5), we get

$${}^{l\downarrow}(Z = \overset{k\rightarrow}{U}(k, l)) = {}^{l\downarrow}\overset{n\rightarrow}{(R)} \cdot {}^{n\downarrow}(X = \overset{m\rightarrow}{u}(m, n)) \cdot {}^{m\downarrow}\overset{k\rightarrow}{(R')}. \tag{7}$$

Comparing the summation in Eq. (5) and the matrix operation in Eq. (7), we get

$$R' = \begin{matrix} \\ m\downarrow \end{matrix}\begin{pmatrix} e^{-i2\pi\frac{0}{N}0} & \cdots & \overset{k\rightarrow}{e^{-i2\pi\frac{k}{N}0}} & \cdots & e^{-i2\pi\frac{N-1}{N}0} \\ \vdots & \ddots & & & \\ e^{-i2\pi\frac{0}{N}m} & & e^{-i2\pi\frac{k}{N}m} & & \\ \vdots & & & \ddots & \\ e^{-i2\pi\frac{0}{N}(N-1)} & & & & e^{-i2\pi\frac{N-1}{N}(N-1)} \end{pmatrix}$$

and

$$R = \begin{matrix} \\ l\downarrow \end{matrix}\begin{pmatrix} e^{-i2\pi\frac{0}{N}0} & \cdots & \overset{n\rightarrow}{e^{-i2\pi\frac{0}{N}n}} & \cdots & e^{-i2\pi\frac{0}{N}(N-1)} \\ \vdots & \ddots & & & \\ e^{-i2\pi\frac{l}{N}0} & & e^{-i2\pi\frac{l}{N}n} & & \\ \vdots & & & \ddots & \\ e^{-i2\pi\frac{N-1}{N}0} & & & & e^{-i2\pi\frac{N-1}{N}(N-1)} \end{pmatrix}. \tag{8}$$

Since these matrices are symmetric, setting

$$W_N = \exp(-\frac{i2\pi}{N}) \tag{9}$$

and

$$R = \begin{pmatrix} W_N^{0\cdot 0} & \cdots & W_N^{0\cdot n} & \cdots & W_N^{0\cdot(N-1)} \\ \vdots & \ddots & & & \\ W_N^{l\cdot 0} & & W_N^{ln} & & \\ \vdots & & & \ddots & \\ W_N^{(N-1)\cdot 0} & & & & W_N^{(N-1)(N-1)} \end{pmatrix}, \tag{10}$$

we get

$$Z = RXR. \tag{11}$$

Now we consider the product of the matrix R and the matrix $R'^{*}(= R^{*})$. Calculating the inner product of the nth column and the complex conjugate of the n'th column, we get

$$\begin{aligned} \sum_{l=0}^{N-1} W_N^{ln} \cdot (W_N^{ln'})^{*} &= \sum_{l=0}^{N-1} \exp(-\frac{i2\pi ln}{N}) \exp(+\frac{i2\pi ln'}{N}) \\ &= \sum_{l=0}^{N-1} \exp(-\frac{i\{(n-n')2\pi\}l}{N}) \\ &= \sum_{l=0}^{N-1} W_N^{(n-n')l}. \end{aligned} \tag{12}$$

This sum is, in case $n \neq n'$,

$$\begin{aligned} \sum_{l=0}^{N-1} W_N^{(n-n')l} &= \frac{1 - W_N^{(n-n')N}}{1 - W_N^{(n-n')}} \\ &= \frac{1 - \left(W_N^{N}\right)^{(n-n')}}{1 - W_N^{(n-n')}} \\ &= \frac{1 - 1^{(n-n')}}{1 - W_N^{(n-n')}} = 0, \end{aligned} \tag{13}$$

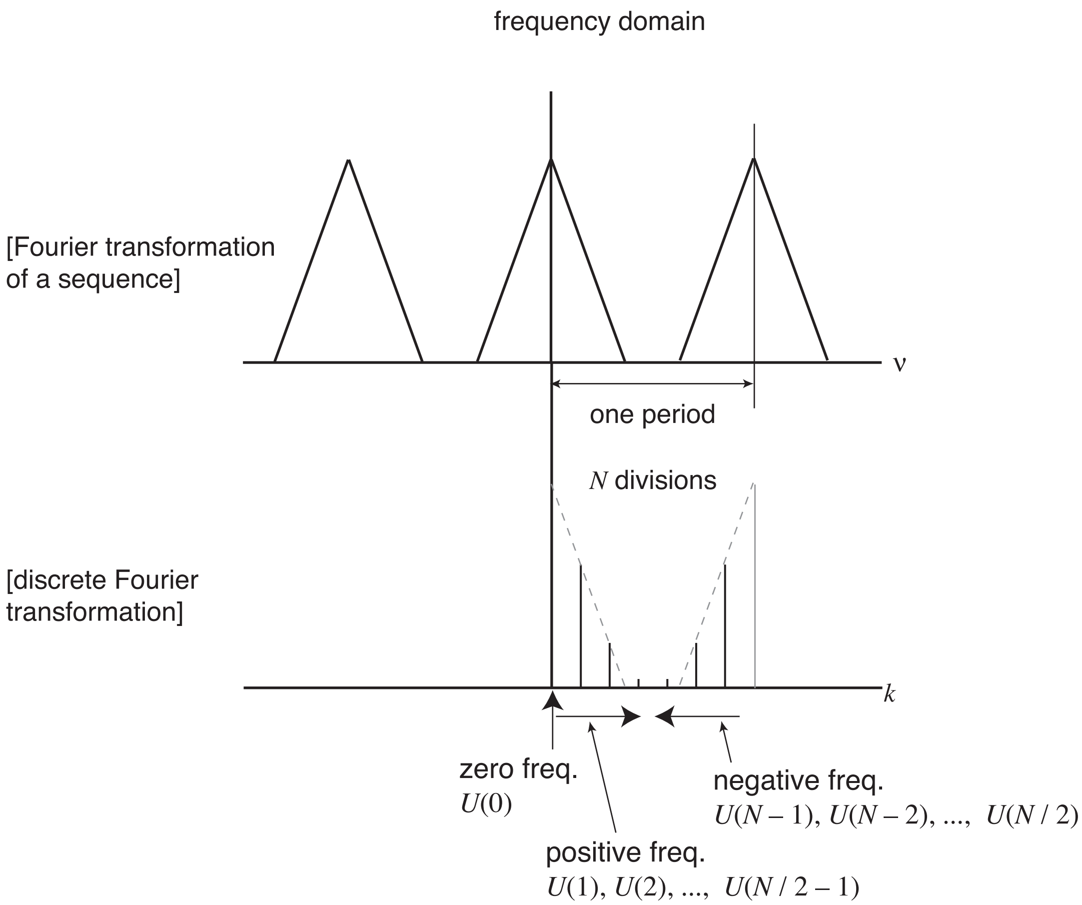

Fig. 1: Correspondance of elements in the discrete Fourier transformation to frequency components.

and in case $n = n'$,

$$\sum_{l=0}^{N-1} W_N^{(n-n')l} = \sum_{l=0}^{N-1} 1 = N. \tag{14}$$

It follows that $RR'^* = NI$, i. e. R is not a unitary matrix. Since RR'^* is the unit matrix multiplied by N, redefining W_N as

$$W_N = \frac{1}{\sqrt{N}} \exp(-\frac{i2\pi}{N}) \tag{15}$$

makes R unitary. In this case, the inverse transformation is as follows:

$$X = R^* Z R^*. \tag{16}$$

The discrete Fourier transformation defined as the above is called *unitary discrete Fourier transformation.*

As shown in the previous topic, sampling the original function causes in the frequency domain a repetition of the frequency distribution of the original function. The discrete Fourier transformation is equivalent to the extraction of the first period of the original frequency distribution from the origin and sampling N points. Thus in the case of the one-dimensional discrete Fourier transformation $U(0)$ corresponds to the component at the frequency zero of the original frequency distribution, $U(1), U(2), \ldots, U(N/2-1)$ correspond to the components at positive frequencies, and $U(N-1), U(N-2), \ldots, U(N/2)$ correspond to the portions at negative frequencies

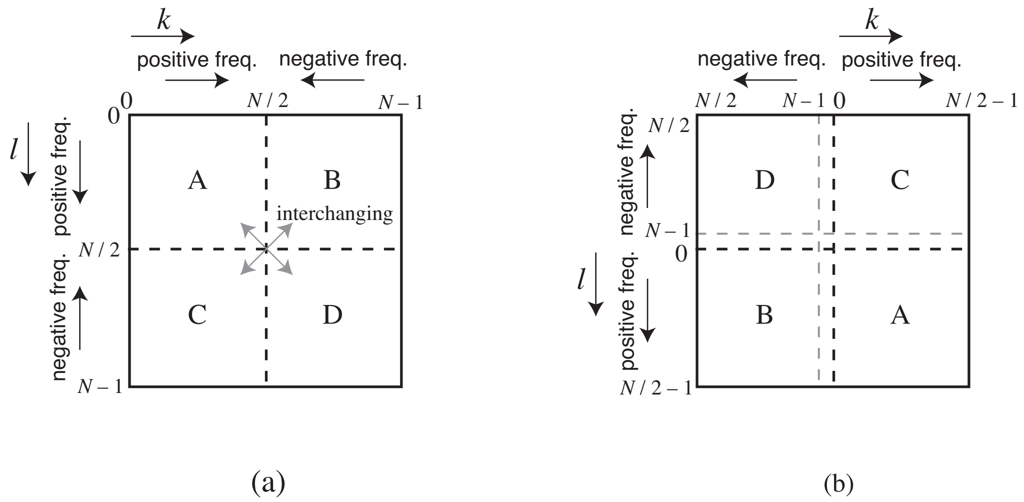

Fig. 2: Correspondance of elements in the twodimensional discrete Fourier transformation to frequency components.

in the reverse order, as shown in Fig. 1. It also holds that the amplitude of Fourier transform of a real function is an even function and the phase is odd. These correspond to the fact that

$$U^*(N-k) = U(k) \tag{17}$$

holds for the one-dimensional Fourier transformation of a real sequence defined in Eq. (3), as explained in Session 3.

Since this holds for both the row operations and the column operations in the two-dimensional case, the elements of the two-dimensional Fourier transform correspond to the frequency components as shown in Fig. 2(a). The interchanges of domains as shown in Fig. 2(b) are convienient for intuitive comprehension.

Since

$$U(k,l) = U^*(M-k, N-l) \tag{18}$$

holds for two-dimensional discrete Fourier transformation of a real matrix similarly to Eq. (17), determining the values in a half of the domain determines automatically the other half, also in the case of the two-dimensional discrete Fourier transformation.

Discrete cosine transformation and image compression

Generally speaking, low-frequency components of an image correspond to brief outlines of image objects, and high-frequency components correspond to fine structures. Thus high-frequency components are usually small and omitting the components causes little loss of image information. This suggests that omitting high-frequency components achieves image compression, as shown at the section of the KL transformation in the Session 4.

The Fourier transformation, however, requires to handle complex numbers. This makes the calculations complicated. Instead of this, *discrete cosine transformation* (DCT), is often used for practical image compression schemes. The matrix R is defined as

$$R = \overset{n\rightarrow}{l\downarrow\begin{pmatrix} \ddots & & \\ & r(n,l) & \\ & & \ddots \end{pmatrix}},$$
$$r(n,l) = \begin{cases} \frac{1}{\sqrt{N}} & l = 0 \\ \frac{2}{\sqrt{N}}\cos\frac{(2n+1)l\pi}{2N} & l \neq 0 \end{cases} \tag{19}$$

instead of Eq. (10) in case of DCT. The matrix R is real in this case. DCT of an image is equivalent to the discrete Fourier transformation of the two-dimensional even function composed by the original image and the reflections with respect to x-axis, y-axis, and the origin [1]. Let us consider the one-dimensional case; Let the original onedimensional signal of N elements be $u(0), u(1), \ldots, u(N-1)$. Connecting this signal with its reflection, we get the sequence of $2N$ elements, $u(N-1), u(N-2), \ldots, u(1), u(0), u(0), u(1), \ldots, u(N-1)$. It follows from Eq. (19) that DCT of a one-dimensional function of N elements is

$$U(k) = \begin{cases} \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n) & k = 0 \\ \frac{2}{\sqrt{N}} u(n)\cos\frac{(2n+1)k\pi}{2N} & k \neq 0 \end{cases}. \tag{20}$$

In case that $k \neq 0$, we get

$$\begin{aligned} U(k) &= \frac{2}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)\cos\frac{(2n+1)k\pi}{2N} \\ &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)\exp\frac{i(2n+1)k\pi}{2N} + \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)\exp\frac{-i(2n+1)k\pi}{2N} \\ &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)\exp\frac{-i((-n)-1/2)k\pi}{2N} + \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)\exp\frac{-i(n+1/2)k\pi}{2N} \end{aligned} \tag{21}$$

This operation is equivalent to the discrete Fourier transformation of the sequence $u(N-1), u(N-2), \ldots, u(1), u(0), u(0), u(1), \ldots, u(N-1)$ by supposing an element between two $u(0)$'s and renumbering the indices of the elements $u(1), u(2), \ldots$ following $u(0)$ to $1/2, 3/2, \ldots$, and the indices of the elements preceding $u(0)$ to $\ldots, -3/2, -1/2$. Since this is the Fourier transformation of the sequence of $2N$ elements, $U(0)$ corresponds to the zero frequency, and the larger k is, the higher frequency component $U(k)$ corresponds to. Since DCT yields a real $N \times N$ matrix from a real $N \times N$ matrix, and larger k corresponds to a higher frequency, it requires no reflection of the transform as in the case of the discrete Fourier transformation. Thus the image compression by omitting high-frequency components is achieved by preserving $U(k,l)$ for smaller k and l and omtting the others. The image compression by the JPEG standard, which is popularly used for practical image compression, is achieved by segmenting an image to domains of 8×8-pixels, transforming each domains by DCT, and omitting higher frequency components of each domain.

[1] The Fourier transform of an even function is a real function (this operation is refered as the Fourer cosine transformation).