

数理科学特論 A ~ 画像数学 ~ 第 6 回

シリーズ 2 : 直交変換による画像圧縮

(3) 離散フーリエ変換と離散コサイン変換

前回は、画像を行列で表し、縦方向と横方向それぞれに作用する直交変換・ユニタリー変換を行う枠組みを示しました。今日は、具体的な直交変換・ユニタリー変換の例と、それによる情報圧縮の方法を説明し、このシリーズを終えたいと思います。

2次元離散フーリエ変換

この講義の第 2 回で、2次元の関数 $f(x, y)$ のフーリエ変換を、

$$F(v_x, v_y) = \iint_{-\infty}^{\infty} f(x, y) \exp\{-i2\pi(v_x x + v_y y)\} dx dy \quad (1)$$

と定義しました。この式は

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

と書き直すことができるので、 x, y それぞれの方向について 1次元のフーリエ変換を行っていることに相当します。また、1次元の N 項の数列 $u(n)$ の離散フーリエ変換は

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

となることも第 3 回の講義で説明しました。これらのことから、 m 方向に M 項、 n 方向に N 項の 2次元数列 $u(m, n)$ の 2次元離散フーリエ変換は

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

$(k=0, 1, \dots, M-1, l=0, 1, \dots, N-1)$ と定義することができます。とくに $M=N$ の場合、つまり正方形のデジタル画像のフーリエ変換を想定すると

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

となります。

さて、前回の講義で、行列 X の縦横に分離可能なユニタリー変換は、ユニタリー行列 R を使って

$$Z = RXR' \quad (6)$$

と表せることを示しました。この形式で (5) 式の変換を表すことを考えてみましょう。(6) 式の行・列と (5) 式の各変数を対応させると

$$\downarrow^l (Z = U(k, l)) = \downarrow^l (R) \cdot \downarrow^n (X = u(m, n)) \cdot \downarrow^m (R') \quad (7)$$

という形になります．そこで，(5) 式の の計算と (7) 式の行列の演算を対応させると，

$$R' = \begin{pmatrix} \downarrow^m & & & & \\ & \overset{k \rightarrow}{e^{-i2\pi \frac{0}{N} 0}} & \dots & e^{-i2\pi \frac{k}{N} 0} & \dots & e^{-i2\pi \frac{N-1}{N} 0} \\ & \vdots & \ddots & & & \\ & e^{-i2\pi \frac{0}{N} m} & & e^{-i2\pi \frac{k}{N} m} & & \\ & \vdots & & & \ddots & \\ & e^{-i2\pi \frac{0}{N} (N-1)} & & & & e^{-i2\pi \frac{N-1}{N} (N-1)} \end{pmatrix} \quad (8)$$

$$R = \begin{pmatrix} \downarrow^l & & & & \\ & \overset{n \rightarrow}{e^{-i2\pi \frac{0}{N} 0}} & \dots & e^{-i2\pi \frac{0}{N} n} & \dots & e^{-i2\pi \frac{0}{N} (N-1)} \\ & \vdots & \ddots & & & \\ & e^{-i2\pi \frac{l}{N} 0} & & e^{-i2\pi \frac{l}{N} n} & & \\ & \vdots & & & \ddots & \\ & e^{-i2\pi \frac{l}{N} (N-1)} & & & & e^{-i2\pi \frac{l}{N} (N-1)} \end{pmatrix} \quad (9)$$

となります．これらは対称行列ですから，

$$W_N = \exp\left(-\frac{i2\pi}{N}\right) \quad (10)$$

とにおいて

$$R = \begin{pmatrix} W_N^{0 \cdot 0} & \dots & W_N^{0 \cdot n} & \dots & W_N^{0 \cdot (N-1)} \\ \vdots & \ddots & & & \\ W_N^{l \cdot 0} & & W_N^{ln} & & \\ \vdots & & & \ddots & \\ W_N^{(N-1) \cdot 0} & & & & W_N^{(N-1)(N-1)} \end{pmatrix} \quad (11)$$

とすると，

$$Z = RXR \quad (12)$$

と表すことができます．

ここで行列 R と行列 R'^* の積を求めてみましょう．行列 R の第 n 列と，第 n' 列の複素共役の内積を求めると

$$\begin{aligned} \sum_{l=0}^{N-1} W_N^{ln} \cdot (W_N^{ln'})^* &= \sum_{l=0}^{N-1} \exp\left(-\frac{i2\pi ln}{N}\right) \cdot \exp\left(\frac{i2\pi ln'}{N}\right) \\ &= \sum_{l=0}^{N-1} \exp\left(-\frac{i\{(n-n')2\pi\}l}{N}\right) = \sum_{l=0}^{N-1} W_N^{(n-n')l} \end{aligned} \quad (13)$$

となります．この値は， $n \neq n'$ のとき

$$\begin{aligned} \sum_{l=0}^{N-1} W_N^{(n-n')l} &= \frac{1 - W_N^{(n-n')N}}{1 - W_N^{(n-n')}} \\ &= \frac{1 - (W_N^N)^{(n-n')}}{1 - W_N^{(n-n')}} = \frac{1 - 1^{(n-n')}}{1 - W_N^{(n-n')}} = 0 \end{aligned} \quad (14)$$

で, $n = n'$ のときは

$$\sum_{l=0}^{N-1} W_N^{(n-n')l} = \sum_{l=0}^{N-1} 1 = N \quad (15)$$

となります。したがって $RR^* = NI$ となり, R はユニタリー行列ではありません。 RR^* は単位行列の N 倍になっているわけですから, R をユニタリー行列とするには (10) 式をすこし変えて, W_N を

$$W_N = \frac{1}{\sqrt{N}} \exp\left(-\frac{i2\pi}{N}\right) \quad (16)$$

と定義すればよいこととなります。このとき, 逆変換は

$$Z = R^*XR \quad (17)$$

となります。このように定義した離散フーリエ変換をとくにユニタリー離散フーリエ変換 (unitary DFT) ということもあります。

さて, もとの関数をサンプリングすると, 周波数空間ではもとの関数の周波数分布が周期的に繰り返されることとなります。離散フーリエ変換は, もとの周波数分布の周波数 0 のところから 1 周期分を取り出して N 等分にサンプリングしたものに相当しています (第 1 シリーズのレポート課題解答例の図 1 を参照してください)。したがって, 1 次元の離散フーリエ変換においては, $U(0)$ が元の周波数分布の周波数 0 に, $U(1), U(2), \dots, U(N/2 - 1)$ まだが正の周波数に, $U(N - 1), U(N - 2), \dots, U(N/2)$ の $N/2$ 個は逆順に負の周波数に相当します (図 1)。また, 実関数のフーリエ変換は振幅が偶関数, 位相が奇関数となります (第 1 シリーズのレポート課題参照)。このことは, 第 3 回に説明したとおり, (3) 式で定義される実数列の 1 次元の離散フーリエ変換について,

$$U^*(N - k) = U(k) \quad (18)$$

がなりたつことに対応しています。

2 次元離散フーリエ変換においては, 縦横両方向に同じ原理が成り立ちますから, 2 次元離散フーリエ変換の要素は図 2 (a) のように周波数に対応します。これは, 図 2 (b) のように領域を入れ替えると分かりやすくなります。

また, 実行列の 2 次元離散フーリエ変換においては (18) 式と同様に

$$U(k, l) = U^*(M - k, N - l) \quad (19)$$

の関係がなりたつので, 2 次元離散フーリエ変換においても全体の $1/2$ の領域の値が決まれば他の値は全て決まることとなります (第 2 シリーズのレポート課題参照)。

離散コサイン変換と画像圧縮

画像では一般に, 低い周波数成分は画像中の物体のおおまかな形を表し, 高い周波数成分は物体の周辺の細かな様子を表しています。したがって, 高い周波数成分はその値も小さく, 省略しても画像が伝える情報はそれほど損なわれません。したがって第 4 回の KL 変換のところで示した考え方にしたがうと, 高い周波数成分を省略することで画像圧縮ができます。

しかし, フーリエ変換は複素数を扱う必要があるため, 計算機での取り扱いがいくぶん面倒です。そこで, 実際の画像圧縮でよく用いられているのが, 離散コサイン変換 (discrete cosine transformation, DCT) です。離散コサイン変換では, 行列 R は (11) 式のかわりに

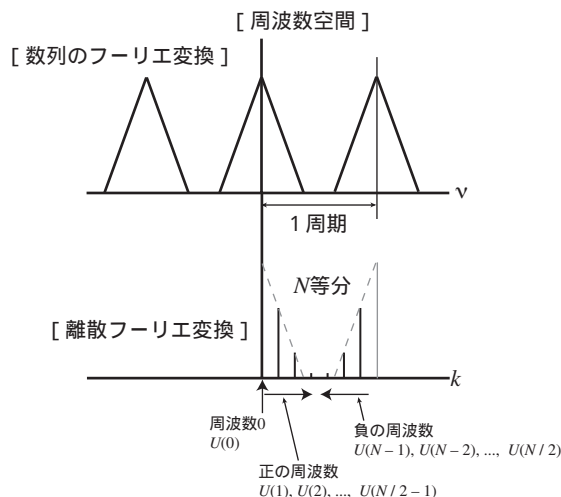


図 1 . 離散フーリエ変換における要素と周波数

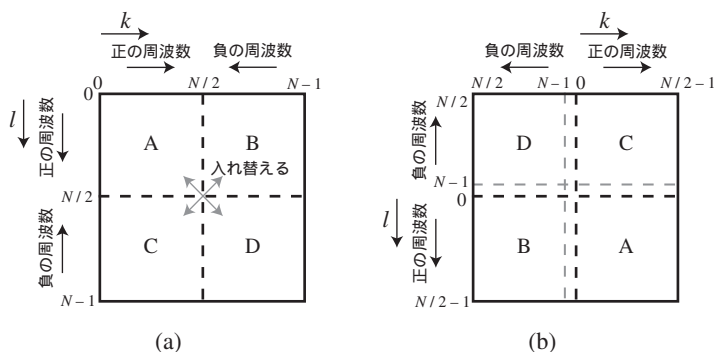


図 2 . 2次元離散フーリエ変換における要素と周波数

$$R = \begin{pmatrix} \ddots & \ddots & \ddots \\ \vdots & r(n, l) & \vdots \\ \vdots & \vdots & \ddots \end{pmatrix}, r(n, l) = \begin{cases} \frac{1}{\sqrt{N}} & l = 0 \\ \frac{2}{\sqrt{N}} \cos \frac{(2n+1)l\pi}{2N} & l \neq 0 \end{cases} \quad (20)$$

と定義されます。離散コサイン変換では、 R は実行列になっています。

離散コサイン変換は、元の画像を縦横とも座標軸に対称に折り返し、偶関数にしたものの離散フーリエ変換に相当します*)。1次元の場合にこれを見てみましょう。元の1次元信号を N 要素の数列 $u(0), u(1), \dots, u(N-1)$ とするとき、これを折り返した数列は $2N$ 要素の $u(N-1), u(N-2), \dots, u(1), u(0), u(0), u(1), \dots, u(N-1)$ となります。 N 要素の1次元の離散コサイン変換は (20) 式から

$$U(k) = \begin{cases} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) & k = 0 \\ \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \cos \frac{(2n+1)k\pi}{2N} & k \neq 0 \end{cases} \quad (21)$$

*) 偶関数の (連続値の) フーリエ変換は実関数になります (フーリエコサイン変換)。

となります。そこで、 $k \neq 0$ のとき

$$\begin{aligned}
 U(k) &= \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \cos \frac{(2n+1)k\pi}{2N} \\
 &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \exp \frac{i(2n+1)k\pi}{2N} + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \exp \frac{-i(2n+1)k\pi}{2N} \\
 &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \exp \frac{-i((-n)-1/2)k\pi}{N} + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(n) \exp \frac{-i(n+1/2)k\pi}{N}
 \end{aligned} \tag{22}$$

となります。これは $u(N-1), u(N-2), \dots, u(1), u(0), u(0), u(1), \dots, u(N-1)$ という数列に対して、折り返し点である 2つの $u(0)$ の間に仮に要素があると考えるとこれを 0番と想定し、それより後を 1/2番、3/2番、...、前を -1/2番、-3/2番、...、と番号をつけて離散フーリエ変換を行ったことに相当しています。2N項の数列のフーリエ変換をしているわけですから、 $k=0$ が周波数 0 に相当し、以後 $k=1, \dots, N \div 1$ と番号が大きいくほど高い周波数に対応する成分となります。

したがって、離散コサイン変換では、 $N \times N$ 行列を変換すると $N \times N$ の実行列が得られ、 $k, l = 0, 1, \dots, N-1$ と番号が大きいくほど高い周波数に対応するので、離散フーリエ変換のように周波数空間で折り返しをする必要はありません。したがって、 k, l とはいくつかの番号までかを限定して残し他を省略することで、高周波成分を省略し情報量を圧縮することができます。画像圧縮でよく用いられている JPEG 方式では、画像を 8×8 画素の領域に分割し、各領域で離散コサイン変換を行って、それぞれの領域で高周波成分を省くことで情報圧縮を行っています。