

# 画像処理実験

## - DCT 符号化による画像圧縮 -

### 1 DCT 符号化による画像圧縮手法

人間の視覚は、画像の高周波成分に鈍感であるという特性を持つ。そこで、画像を高周波成分から低周波成分に周波数に分解し、高周波成分の情報を除外することにより、画質を保ったまま、画像の情報量の削減が期待できる。画像を高周波成分から低周波成分に分解する手法として、離散コサイン変換 (DCT: Discrete Cosine Transform) が主に用いられている。1 次元の  $DCT(F(u))$  とその逆変換 ( $f(x)$ ) (IDCT: Inverse DCT) は、次式で表される。

$$F(u) = \frac{2}{N} C(u) \sum_{x=0}^{N-1} f(x) \cos \frac{(2x+1)u\pi}{2N} \quad (1)$$

$$f(x) = \frac{2}{N} \sum_{u=0}^{N-1} C(u) F(u) \cos \frac{(2x+1)u\pi}{2N} \quad (2)$$

ただし、

$$u = 0 \text{ のとき } C(u) = \frac{1}{\sqrt{2}} \\ u \neq 0 \text{ のとき } C(u) = 1$$

とする。ただし、 $N$  は、信号のダイナミックレンジを表す。ここで、画像に対して DCT 変換するために、2 次元の DCT 変換を考える。2 次元の  $DCT(F(u, v))$  とその逆変換 ( $f(x, y)$ ) は、次式で表される。

$$F(u, v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (3)$$

$$f(x, y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u, v) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (4)$$

ただし、

$$u = 0, \text{ または }, v = 0 \text{ のとき } C(u) = \frac{1}{\sqrt{2}}, \text{ または }, C(v) = \frac{1}{\sqrt{2}} \\ u \neq 0, \text{ または }, v \neq 0 \text{ のとき } C(u) = 1, \text{ または }, C(v) = 1$$

とする。

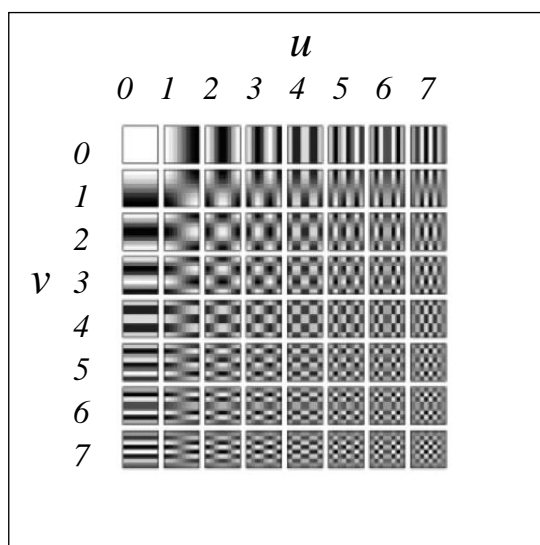


図 1: DCT で用いる基底の画像

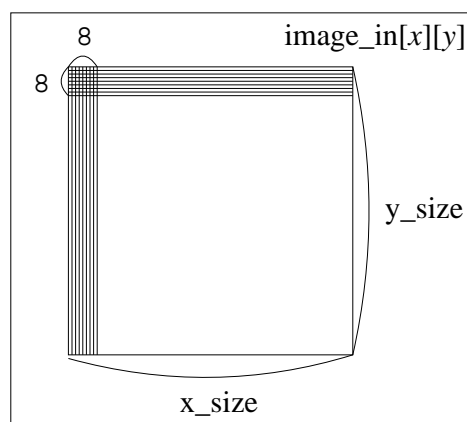


図 2: 画像の輝度値を格納する image\_in 配列

## 2 実験

### 2.1 実験 1 (DCT 符号化の実装)

画像入出力プログラム (imageio.c) を改造して、与えられた画像を DCT 符号化し、情報量を削減した画像を出力するプログラムを実装せよ。ただし、(3)、(4) 式の  $N$  を 8 として実装せよ。

### 2.2 実験 1 における DCT 符号化の実装方法

画像の輝度値を格納する配列を  $\text{image\_in}[x][y]$  とする。この配列の大きさは、 $x$  軸方向、 $y$  軸方向それぞれ、 $x\_size, y\_size$  とする (図 2)。また、DCT 係数を格納する配列を  $c[x+u][y+v]$  とする。DCT 係数は、各ブロック  $8 \times 8$  個あり、ブロックの個数は、 $x$  軸方向、 $y$  軸方向それぞれ  $x\_size/8, y\_size/8$  個あるので、 $x$  軸方向の総ブロックサイズは  $x\_size$ 、 $x$  軸方向の総ブロックサ

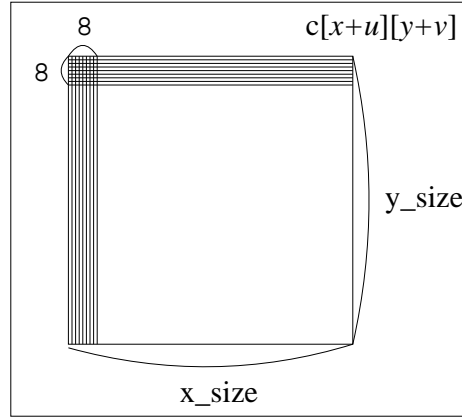


図 3: DCT 係数を格納する c 配列

イズは  $y\_size$  となる (図 3)。

各ブロックにおける係数を求めるには、

$x$  は 0 から  $x\_size-1$  まで、8 画素ごと、増やし、

$y$  は 0 から  $y\_size-1$  まで、8 画素ごと、増やし、

$u$  は 0 から 7 ( $u$  方向の基底数  $-1$ ) まで、1 つずつ、増やし、

$v$  は 0 から 7 ( $v$  方向の基底数  $-1$ ) まで、1 つずつ、増やして、

式 (3) に基づき DCT 係数を以下のように計算し、 $c[x+u][y+v]$  に格納する。

$$c[x+u][y+v] = \frac{2}{N} C(u) C(v) \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \text{image\_in}[x+j][y+k] \cos \frac{(2j+1)u\pi}{2N} \cos \frac{(2k+1)v\pi}{2N} \quad (5)$$

各ブロックにおいて求めた係数から情報量を削減した画像を得るには、

$x$  は 0 から  $x\_size-1$  まで、8 画素ごと、増やし、

$y$  は 0 から  $y\_size-1$  まで、8 画素ごと、増やし、

$j$  は 0 から 7 (ブロックサイズ  $-1$ ) まで、1 つずつ、増やし、

$k$  は 0 から 7 (ブロックサイズ  $-1$ ) まで、1 つずつ、増やして、

式 (4) に基づき画像の輝度値を計算し、 $\text{image\_out}[x+j][y+k]$  に格納する。

$$\text{image\_out}[x+j][y+k] = \frac{2}{N} \sum_{u=0}^{BL-1} \sum_{v=0}^{BL-1} C(u) C(v) c[x+u][y+v] \cos \frac{(2j+1)u\pi}{2N} \cos \frac{(2k+1)v\pi}{2N} \quad (6)$$

ただし、 $BL$  は帯域制限数 (逆変換に用いる基底数、1 から 8 まで) とする。

## 2.3 実験 2 (情報量削減と画質の関係)

$PSNR$  を画質の尺度とし、高周波成分を除外した量と画質の関係をグラフにまとめ、考察せよ。なお、 $PSNR$  は、次式で表される。

$$PSNR[dB] = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad (7)$$

ただし、

$$MSE = \frac{1}{AB} \sum_{x=1}^A \sum_{y=1}^B |I_o(x, y) - I_c(x, y)|^2$$

とし、 $A$  は、画像の横方向のサイズ、 $B$  は画像の縦方向のサイズ、 $I_o(x, y)$  は、原画像の画素  $(x, y)$  における輝度値を表し、 $I_c(x, y)$  は、情報削減した画像の画素  $(x, y)$  における輝度値を表す。

画像入出力プログラム (imageio2.c) を改造して、原画像と情報量を削減した画像との  $PSNR$  を算出するプログラムを実装せよ。帯域制限 ( $BL$ ) の変化に対する  $PSNR$  の変化をグラフにまとめ、関係を考察せよ。

## 2.4 実験 2 における情報量削減と画質の関係の評価方法

グラフは、gnuplot というアプリケーションを使って、作成する。まず、以下の内容のファイル (graph.plt) を作成する

```
set term X11
# set term postscript eps
set nokey
set grid xtics ytics
show grid
set xtics 1
set ytics 5
set size 0.5, 0.5
set xlabel "Band Number"
set ylabel "PSNR[dB]"
#set output "graph.eps"
plot [1:7][20:50] 0 t ' ' w l 0, "graph.dat" t ' ' w linespoint 1 2
pause -1
```

次に、実験データファイル (graph.dat) を作成する。

例

#(横軸データ) (縦軸データ)

1	22.64
2	34.73
3	37.21
4	38.75
5	39.12
6	39.89
7	40.02

グラフを表示するには、コマンドラインで

% gnuplot graph.plt

とする。eps ファイルで出力するには、graph.plt の#がついている部分をはずす。作成したグラフの特性を考察し、まとめる。

### 3 考察

実験結果に基づき、以下の項目について考察せよ。

1.  $BL$  を変化させたときの  $PSNR$  と人間の主観的な画像評価は、一致するか。実験で作図したグラフと、画像圧縮した画像を示し、考察せよ。
2.  $BL$  を変化させたときどの程度まで人間は画質劣化が認識できないか。そのときの  $PSNR$  の値はどうか。また、その際の削減可能な情報量はどの程度か。
3. 一般的に信号の評価には信号対雑音比 ( $SNR$ ) が用いられるが、画質評価には、なぜ  $PSNR$  が用いられるのか。
4. 画像圧縮には  $\sin$  変換ではなく、一般的に  $\cos$  変換が用いられているが、なぜか理由を示せ。
5. DCT 変換以外の画像圧縮手法を 2 つあげ、それぞれについて説明せよ。