

## 1.1 画像の評価法

画像の評価方法としては主に原画像と処理画像とを数値で比較する客観評価と、人が見て評価をする主観評価がある。客観評価，主観評価共に幾つかの種類がある。

### 1.1.1 平均二乗誤差 (MSE)

画像の数値評価法として良く使われる方法が平均二乗誤差 (Mean Square Error) ,M S E である。MSE は以下の式で定義される。

$$MSE = \frac{1}{M \times M'} \sum_{i=1}^M \sum_{j=1}^{M'} e(i, j)^2 = \frac{1}{M \times M'} \sum_{i=1}^M \sum_{j=1}^{M'} \{y(i, j) - s(i, j)\}^2$$

M S E は原画像と処理画像との差の 2 乗誤差である。MSE が小さければ小さいほど原画像に近い画像であり，主に復元などの処理の評価に用いられる。また，正規化 MSE (normalized MSE) は以下の式で定義される。

$$NMSE = \frac{\sum_{i=1}^M \sum_{j=1}^{M'} \{y(i, j) - s(i, j)\}^2}{\sum_{i=1}^M \sum_{j=1}^{M'} s(i, j)^2}$$

M S E は画像処理における基本的な評価指標であるが，予め原画像を用意しておく必要があるため，画像処理のフィルタ係数等の設定は理想画像に推定雑音を重畳させた画像で行うことが必要となる。

### 1.1.2 SNR (Signal to Noise Ratio)

S N R は信号レベルに対する雑音レベルの比であり，雑音の大きさ N に対する信号の大きさ S の比で定義され，以下の様に与えられる。

[無単位]

$$SNR = \frac{S}{N}$$

[db]

$$SNR = 10 \log_{10} \frac{S}{N}$$

SNR の数値が大きいと雑音レベルが少なく，画質としては良好である。よって，画像にどれだけ雑音が重畳しているかを知るために使われることが多い。

信号レベル(S)としては画像信号値の平均 2 乗や最大値  $P_s$  が主に使われ，雑音レベル(N)としては MSE 等が用いられる。人間の感覚の強弱はウェバー・フェヒナーの法則から刺激量の対数に

比例するため、感覚的な評価量として扱う場合にはデシベル[db]で算出した方がよいと言われている。

例えば SNR [db]は

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^M \sum_{j=1}^{M'} y(i, j)^2}{\sum_{i=1}^M \sum_{j=1}^{M'} \{y(i, j) - s(i, j)\}^2}$$

定義でき、また最大値  $P_s$  を用いたときには PSNR (Peak SNR) と呼ばれる。

### 1.1.3 詳細分散と背景分散

MSE は原画像との比較による評価指標であり、強調された原画像がない MSE では強調の度合いを評価することができない。よって、雑音画像の強調における評価指標として、詳細分散 (Detail Variance : DV) 及び背景分散 BV (Background Variance : BV) がある<sup>32)</sup>。このとき DV 及び BV は以下の式で示される。

$$DV = \frac{1}{M} \sum_{(i,j) \in A} \{x(i, j) - xd\}^2$$

$$xd = \frac{1}{M} \sum_{(i,j) \in A} x(i, j)$$

$$BV = \frac{1}{K-M} \sum_{(i,j) \in B} \{x(i, j) - xb\}^2$$

$$xb = \frac{1}{K-M} \sum_{(i,j) \in B} x(i, j)$$

領域Aの信号の分散が DV で、領域Bの信号の分散が BV である。領域Aは局所分散がしきい値  $th$  以上の画素から構成される画像のエッジや細部にあたる領域である。逆に領域Bは局所分散がしきい値  $th$  より小さい画素から構成される画像の平坦部分に対応する領域である。また、Kは画像全体の画素数であり、Mは領域Aの画素数である。しきい値  $th$  は画素毎に局所分散を用いてヒストグラムを作成し、そのヒストグラム上で2値化処理を実行したときに得られるしきい値である<sup>55)</sup>。

理想的な強調が行われた場合、画像のエッジのみが強調されることから、処理前の画像と比較してDVの値は大きな値を示す。一方、雑音は強調されないことから、平坦の分散を示すBVの値は変化しないことになる。よって、DVが大きくBVが小さければ理想的な強調がなされたと考えられることができる。

### 1.1.4 色再現の評価方法

人間の視覚が感じ取れる知覚誤差は色の三原色である RGB での誤差とは異なるため、知覚的な誤差を測定するために、国際照明委員会 (CIE) が知覚的に均等な尺度を持つ色空間として L\*a\*b\* 色空間 (CIE LUV) が定義されている<sup>51)</sup>。そこで L\*a\*b\* 色空間での原画像と出力画像との誤差を比較することで、知覚的な誤差評価に近い評価が可能であり、その評価式 NCD (normalized Color Difference) は以下のように定義される。

$$NCD = \frac{\sum_{i=0}^{N_1} \sum_{j=1}^{N_2} \|\Delta E_{L_{ab}}\|}{\sum_{i=0}^{N_1} \sum_{j=1}^{N_2} \|E_{L_{ab}}^*\|}$$

$$\Delta E_{L_{ab}} = \sqrt{(L(i,j) - L_s(i,j))^2 + (a(i,j) - a_s(i,j))^2 + (b(i,j) - b_s(i,j))^2}$$

$$E_{L_{ab}}^* = \sqrt{L_s(i,j)^2 + a_s(i,j)^2 + b_s(i,j)^2}$$

なお、L\*a\*b\* 空間での L は明度であり、a 及び b は彩度となる。また、出力画像は RGB 色空間のままであるので、L\*a\*b\* 色空間への変換が必要となる<sup>52)</sup>。この変換手順は次のように行う。RGB 空間を NTSC 方式に基づいた XYZ 空間へ変換し、その後 XYZ 空間を L\*a\*b\* 色空間へ変換する

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & \text{if } Y/Y_n > 0.008856 \\ 903.29(Y/Y_n) & \text{if } Y/Y_n \leq 0.008856 \end{cases}$$

$$a^* = 500 \cdot [f(X/X_n) - f(Y/Y_n)]$$

$$b^* = 200 \cdot [f(Y/Y_n) - f(Z/Z_n)]$$

$$f(x) = \begin{cases} x^{1/3} & \text{if } x > 0.008856 \\ 7.78x + 16/116 & \text{if } x \leq 0.008856 \end{cases}$$

また、 $x_n, y_n, z_n$  は NTSC 方式では

$$[x_n \ Y_n \ Z_n] = [98.072 \ 100.000 \ 118.225]$$

が用いられる。これは完全拡散型の標準の光による色度である。

### 1.1.5 主観評価

主観評価方法については動画やTV画像で使われる主観評価と同じ評価方法が用いられる。主観評価の方法には対象とする画像のみで評価する場合と複数の画像を同時に比較して評価する場合がある<sup>1,53)</sup>。

単一で評価する場合には、順序付けられたカテゴリに数値を与え、複数の評価結果から数値的な評点を得る評定尺度法を用いることが多い。例えば、「非常に良い」から「非常に悪い」までの間を「良い」「普通」「悪い」のような複数のカテゴリに分ける。この場合は5つのカテゴリに分けられ、それぞれに対して最高点を5点、最低点を1点というように与え、複数の評定者から結果を得る。得られた結果は以下の平均オピニオン評定(MOS, Mean Opinion Score)によって算出する。

$$MOS = \frac{1}{n} \sum_{t=1}^n \alpha_t$$

ここで、 $n$ は評定者数、 $i$ は評定者 $i$ の評価値である。このMOSは非常に簡易的な方法であるので一般的に広く用いられるが、評定者の数を多くする必要があり、15名から20名程度の人数でMOSの標準偏差が0.1程度までになることが報告されている<sup>1)</sup>。

この方法の具体的な例として、複数の方法で作成した画像をディスプレイ上にランダムに表示させ、それぞれの画像の作成方法は評定者に伏せた状態で評価を行ってもらい、その結果をMOSによって算出するという手順となる<sup>54)</sup>。

複数の画像を同時に比較して評価する場合には、例えば「まったく差が無い」から「非常にさがある」までの間を「わずかに差がある」「差がある」「かなり差がある」のような複数のカテゴリに分ける一対比較が用いられることが多い。しかしながら、評価すべき画像や項目が多数ある場合にはその組合せは膨大な数となるため、多変量解析を用いて解析することになる。これについての詳細は文献53を参照されたい。