

■2 群 (画像・音・言語) -5 編 (画像符号化)

1 章 画像・映像信号の性質

(執筆者：吉田俊之) [2013 年 4 月 受領]

■概要■

対象物体の像を記録したものを画像あるいは映像と呼び、これを信号として捉えて画像・映像信号と呼ぶ。アナログ/デジタル画像、カラー/モノクロ画像、静止/動画像などに分類される。画像符号化が対象とするデジタル画像は、物体像を空間サンプリングして画素の配列として表現するとともに、各画素値を量子化することでデジタル表示を可能としている。画像、特に動画像信号のもつ情報量は膨大であるため、高い圧縮効率を有する符号化技術が不可欠となる。

一般に、画像信号は高い画素間相関を有する。これは、自己相関関数あるいは自己共分散関数を用いて定量化され、後者から導かれる画素間相関係数が極めて 1 に近い値を呈するという性質に帰着される。レート-ひずみ理論によれば、画像信号を圧縮符号化できる根拠は、このような高い画素間相関性に基づいている。

画像符号化の実際の観点からは、動画像におけるインタレースとプログレッシブ走査方式、及び色信号の扱いについて理解しておく必要がある。

【本章の構成】

本章では、画像・映像信号とその基本的性質について、画像・映像信号とは (1-1 節)、画像・映像信号の情報量 (1-2 節)、画像・映像信号の統計的性質 (1-3 節)、インタレースとプログレッシブ走査 (1-4 節)、及び色信号のフォーマット (1-5 節) について述べる。

■2群 - 5編 - 1章

1-1 画像・映像信号とは

(執筆者：吉田俊之) [2013年4月 受領]

レンズを介して結像した対象物体の像を記録したものを画像や映像、あるいは信号として捉えて画像・映像信号と呼ぶ。像各点の輝度のみ、または色を含めて記録したものが、各々モノクロ、カラー画像（信号）である。カラー画像の記録には、三原色の原理に基づいて可視光領域の独立な三つの成分を記録する必要があり、実際の三原色には赤（R）、緑（G）、青（B）が用いられる。

画像の記録方式はアナログとデジタル方式に大別される。銀塩写真は前者の代表例で、像を空間的な連続信号として記録するのに対し、後者は像を適当な間隔で空間サンプリングし、各サンプル点上での輝度／色を記録する。

一般に、記録対象の像各点の輝度／色は時間とともに変化する。その一瞬を記録したものが静止画像、適当な時間間隔で連続的に記録したものが動画像である。デジタル動画像を構成する各点（これを画素、pixel, pel と呼ぶ）の画素値 f （輝度あるいは RGB 各成分を表す）は、水平、垂直方向の位置を表す変数 i, j 、及び時刻に相当する変数 n を用いて

$$f(\Delta_x \cdot i, \Delta_y \cdot j, \Delta_t \cdot n) \quad (i, j, n \text{ は整数}) \quad (1 \cdot 1)$$

と表される。ここで、 $\Delta_x, \Delta_y, \Delta_t$ は各々水平、垂直、時間方向のサンプリング間隔（定数）で、 i, j, n は整数値のみをとり得ることに注意されたい。動画像は、各 n に対応する単一静止画像の集合と捉えることができ、各静止画像をフレーム、 Δ_t の逆数をフレームレート（単位時間当たりのフレーム数に相当）と呼ぶ。

■2群 - 5編 - 1章

1-2 画像・映像信号の情報量

(執筆者：吉田俊之) [2013年4月 受領]

水平方向に L_x , 垂直方向に L_y の広がりをもつ像を式(1・1)に従って空間サンプリングすると, 水平, 垂直方向に各々 $W=L_x/\Delta_x$, $H=L_y/\Delta_y$ の画素が生ずる. 画像のサイズは, 水平 (垂直) 方向の画素数 W (H) で規定し, これをサイズ $W \times H$ [画素] の画像と呼ぶ. 画像サイズとしては, HD (1920×1080), SD (720×480), VGA (640×480), SIF (352×240), CIF (352×288), QCIF (176×144) などのフォーマットが規格化されている.

式(1・1)で表される各画素値は実数であり, デジタル画像信号として記録するため, 及び情報量の削減のために量子化を行う. 通常の画像では, 輝度/RGB 成分とも, 最低-最大値間を 8 ビット (256 階調) に一様量子化する. 高精細画像では 8 ビットを越えるビット数が用いられることもある.

サイズ $W \times H$ [画素], フレームレート T [フレーム/秒], RGB 各 8 ビット量子化のデジタルカラー映像信号の情報量は, 毎秒当たり $24WHT$ [ビット/秒] ($3WHT$ [バイト/秒]) となる. 30 [フレーム/秒] の HD 画像の場合には 186.6 M バイト/秒にも達し, 圧縮符号化が必要となるのはこのためである.

■2群 - 5編 - 1章

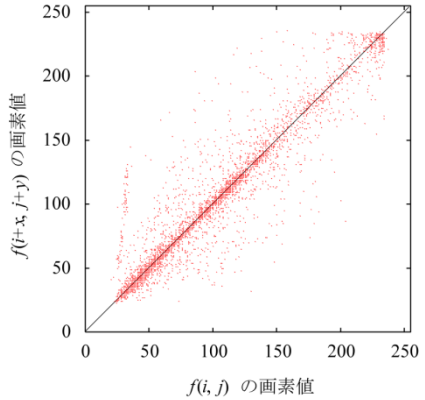
1-3 画像・映像信号の統計的性質

(執筆者：吉田俊之) [2013年4月 受領]

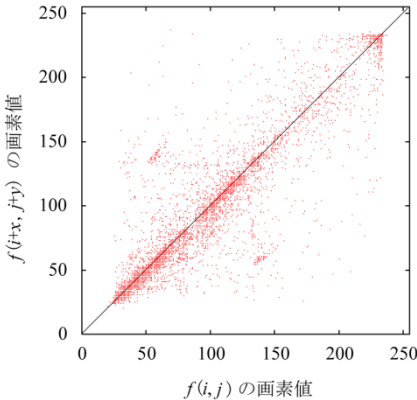
本節では、画像（信号）とは一般的な自然画像を指し、白色雑音を縦横に並べた配列やCGなどの特殊なものは除外して考える。一般に、画像信号には「画素間相関が高い」、すなわち「近接画素の画素値は互いに近い」という性質がある。これは、以下に述べる、(1)散布図が45度の直線に沿って分布する、(2)自己共分散関数が原点からなだらかに減少する、(3)画素間の相関係数が1に近い値をとる、などから説明される。



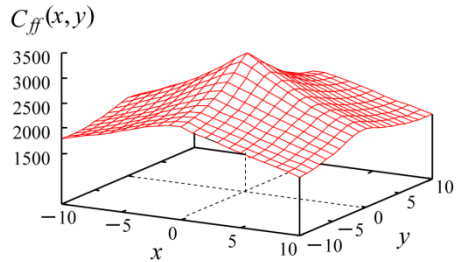
(a) 原画像 $f(i, j)$



(b) $(x, y) = (1, 0)$ の散布図



(c) $(x, y) = (0, 2)$ の散布図



(d) 自己共分散関数 $C_{ff}(x, y)$

図 1・1 画像の画素間相関

図 1・1(b),(c)は、(a)の画素値 $f(i, j)$ に対し、互いに (x, y) だけ離れた画素ペア $(f(i, j), f(i+x, j+y))$ の値をプロットした散布図である。各点は45度の直線の周囲に分布しており、これよ

り水平、垂直方向とも画素間相関が高いことが確認できる¹⁾。

画素間相関は、自己相関関数 (\cdot_{ff} は、 f 自身との相関であることを示す)

$$R_{ff}(x, y) = \frac{1}{N} \sum_i \sum_j f(i, j) f(i+x, j+y) \quad (1 \cdot 2)$$

あるいは、 $f(i, j)$ の空間平均値 \bar{f} を明示的に除いた自己共分散関数

$$C_{ff}(x, y) = \frac{1}{N} \sum_i \sum_j (f(i, j) - \bar{f})(f(i+x, j+y) - \bar{f}) \quad (1 \cdot 3)$$

を用いて定量化できる (N は総和する総画素数)。両式は、 $f(i, j)$ とそれを (x, y) だけ移動した $f(i+x, j+y)$ との「波形の近さ」を定量化している点に注意。詳細は参考文献 1), 2) など参照)。

図 1・1(d)は(a)に対して x と y を変化させて $C_{ff}(x, y)$ を計算した結果で、 C_{ff} は原点からなだらかに減少していることから、(a)の画素間相関が高いことが確認できる²⁾。 $C_{ff}(0, 0)$ が f の分散 σ^2 に相当することを利用し、 $C_{ff}(x, y)/C_{ff}(0, 0)$ と規格化した値を相関係数 $\rho(x, y)$ と呼ぶ。画像の分野では、特に $\rho(1, 0)$ 、 $\rho(0, 1)$ を各々 ρ_H 、 ρ_V とおき、水平、垂直方向の画素間相関係数と呼ぶ。一般に、画像に対しては ρ_H 、 ρ_V ともに 0.95~0.98 程度の極めて 1 に高い値を呈する³⁾。また、画像の $C_{ff}(x, y)$ は、

$$C_{ff}(x, y) = \sigma^2 \rho_H^{|x|} \rho_V^{|y|} \quad (1 \cdot 4)$$

によって非常に良好に近似できることも知られている。

$C_{ff}(x, y)$ のフーリエ変換 $S_{ff}(\omega_x, \omega_y)$ はパワースペクトル密度関数と呼ばれ、周波数当たりの電力を表す。 $C_{ff}(x, y)$ が式(1・4)のように近似され、極めてゆっくりと変動する場合は、そのフーリエ変換である $S_{ff}(\omega_x, \omega_y)$ は原点 (直流) 付近に集中する。レート-ひずみ理論³⁾によれば、波形 $f(i, j)$ を非可逆符号化する場合のレートとひずみの関係は S_{ff} によって定まることが知られている。 $S_{ff}(\omega_x, \omega_y)$ の集中度が高いほど、一定符号化レートの下でのひずみを低く抑えることが可能で、これが画像を圧縮符号化できる根拠である。また、これは画像領域では画素間相関が高い、あるいは相関係数が極めて 1 に近いという性質に基づくことに注意されたい。

■2群 - 5編 - 1章

1-4 インタレースとプログレッシブ走査

(執筆著：吉田俊之) [2013年4月 受領]

かつてのアナログテレビ放送では、走査によって各フレームを1次元波形に変換してアナログ伝送していた。走査とは、フレームを垂直方向の走査線に分解する操作を指すが、これは本質的に垂直方向のサンプリングである。

一般に、動画像内の動きは、式(1・1)におけるフレーム間隔 Δ_t を小さくとることにより滑らかに表示できるが、同時にフレーム数が増加し、記録・伝送すべき情報が増大してしまう。情報量を抑えつつ動きを滑らかに表示する手法として考え出されたのが飛び越し (Interlace) 走査である。

図1・2(a)は、時間間隔 Δ_t でフレーム内のすべての画素 (あるいは走査線) をサンプリングする方式で、順次 (Progressive) 走査方式と呼ぶ。一方、図1・2(b)は、走査線を奇数番目と偶数番目に分け、各々 (奇フィールド、偶フィールドと呼ぶ) を時間軸に沿って交互にサンプリングする方式で、飛び越し (Interlace) 走査方式と呼ばれる。

順次走査(a)と飛び越し走査(b)を比較すると、単位時間当たりに生成される画素数 (走査線数) は同一である一方、後者は見かけ上、時間方向のサンプリング間隔が半分になっている。すなわち、飛び越し走査は、順次走査と同じ情報量でより滑らかな動きを再生できるため有効な走査方式と考えられ、歴史的経緯から現在のデジタル放送においても利用されている。なお、飛び越し走査は時間-垂直平面内では「五の目状サンプリング」に相当し、サンプリング理論に基づく厳密な帯域制限を施すことは困難であるため、折り返しひずみによる画質の低下が指摘されている⁴⁾。

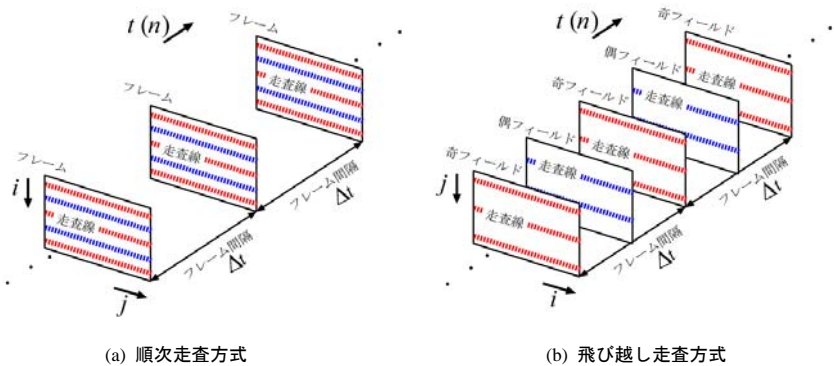


図1・2 順次 (Progressive) 走査と飛び越し (Interlace) 走査

■2群 - 5編 - 1章

1-5 色信号のフォーマット

(執筆著：吉田俊之) [2013年4月 受領]

カラー画像の RGB 各成分間には高い相関があり，高圧縮を実現するためには成分間の相関を除去する必要がある．これには，例えば直交変換などを用いる方法も適用できるが，実際には人間の視覚特性に基づく巧みな手法を利用している．

我々は，色を見る際に「色あい」だけでなく，その「明るさ（輝度）」も感じ取っている．各画素の輝度値を Y で表すと， Y は RGB 各成分 R ， G 及び B から

$$Y = 0.299 R + 0.587 G + 0.114 B \quad (1 \cdot 5)$$

によって求められる⁵⁾．更に，色差成分と呼ばれる

$$C_b = 0.564 (B - Y), \quad C_r = 0.713 (R - Y) \quad (1 \cdot 6)$$

を作ると，各々は「青み成分」，「赤み成分」に対応する．

一方，我々の視覚特性には，“輝度に対する感度は高く，色に対する感度は低い”という性質がある⁵⁾．これを利用すると，画質を維持しながら，輝度 Y に対して色差 C_b ， C_r の解像度（画素数）を落すことができる．図 1・3 に示すように，色差成分 C_b ， C_r の画素数を，(a) に対して水平方向にのみ 1/2 に削減したものを 4:2:2 形式，水平垂直両方向とも 1/2 に削減したものを 4:2:0 形式と呼び，削減しない(a)を 4:4:4 形式と呼ぶ．画像符号化では，対象となる RGB カラー画像を，まず 4:2:0 または 4:2:2 形式に変換した後に， Y ， C_b ， C_r 各成分を独立に符号化することで，符号化効率を上げている．

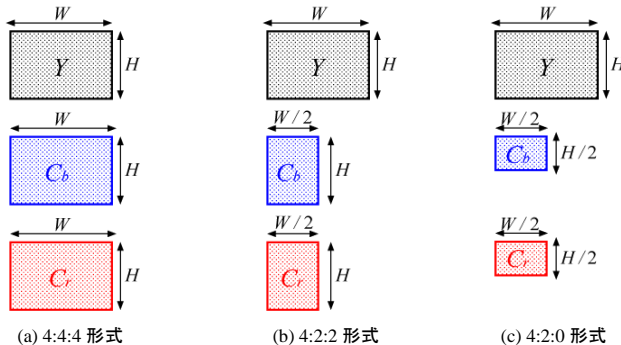


図 1・3 カラー画像のフォーマット

■参考文献

- 1) 長尾 真(監訳)，“デジタル画像処理,” pp.37-46, 近代科学社, 1978.
- 2) 酒井善則, 吉田俊之, “映像情報符号化,” pp.34-38, オーム社, 2001.
- 3) N. S. Jayant and P. Noll, “Digital coding of waveforms,” pp.632-649, Prentice-Hall, Englewood Cliffs, 1984.
- 4) 吹抜敬彦, “テレビ信号の多次元信号処理,” pp.91-95, 日刊工業新聞社, 1988.
- 5) 日下秀夫(監修)，“カラー画像工学,” オーム社, 1997.