

## 第4章 マルチメディアの基礎技術

4-1 マルチメディア情報システムの基本構成

4-2 AD変換

4-2-1 標本化

4-2-2 量子化

4-2-3 符号割り当て

4-3 マルチメディアの情報源符号化技術

4-3-1 情報源符号化の目的

4-3-2 レートとひずみの関係

4-3-3 情報源符号化の要素技術

4-3-4 画像符号化と国際標準化

4-3-5 音声・音響符号と国際標準化

4-4 マルチメディアの通信路符号化技術

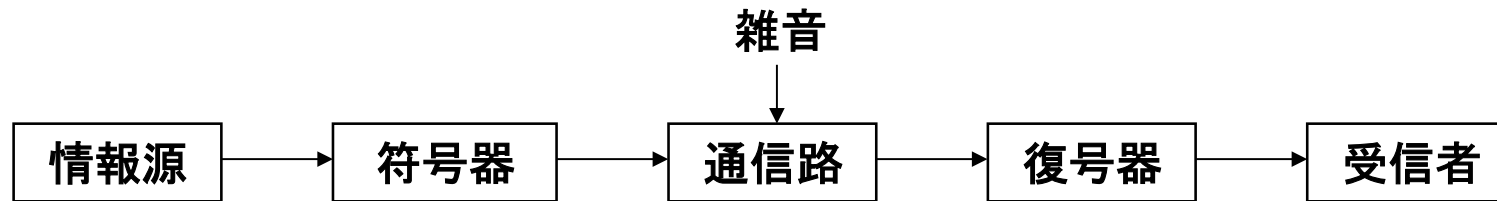
4-4-1 メディア多重化技術

4-4-2 暗号化技術

4-4-3 誤り検出／訂正技術

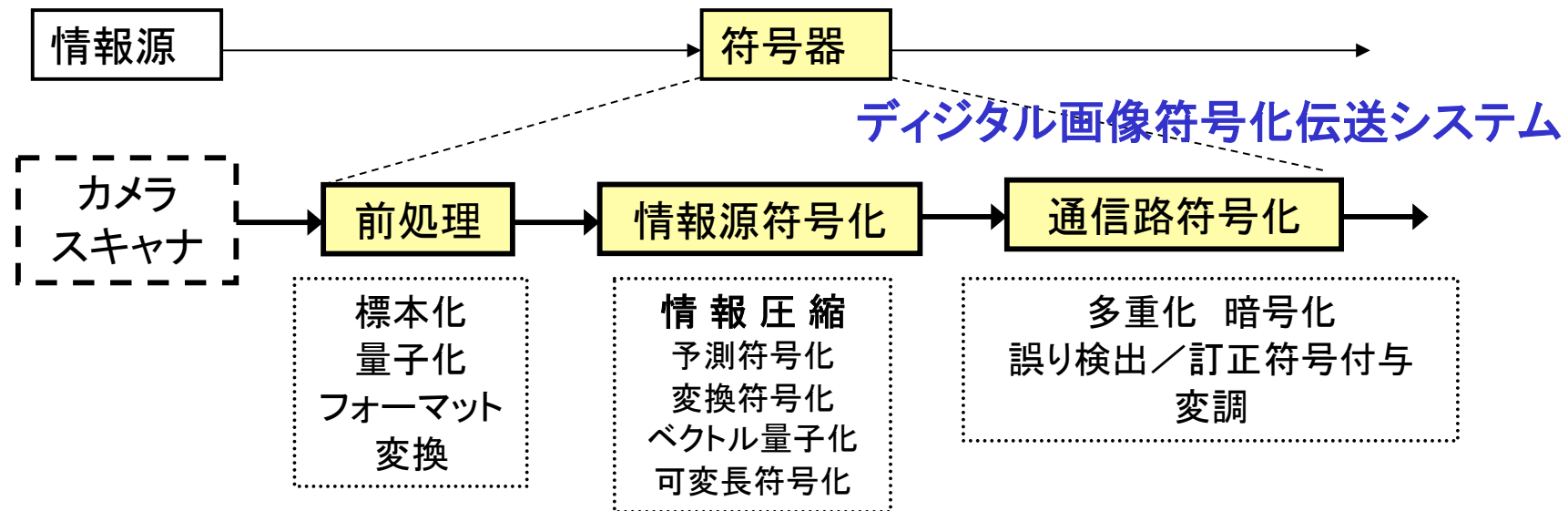
4-4-4 変調技術

## 4-1. マルチメディア情報処理システムの基本構成



シャノンの通信系モデル

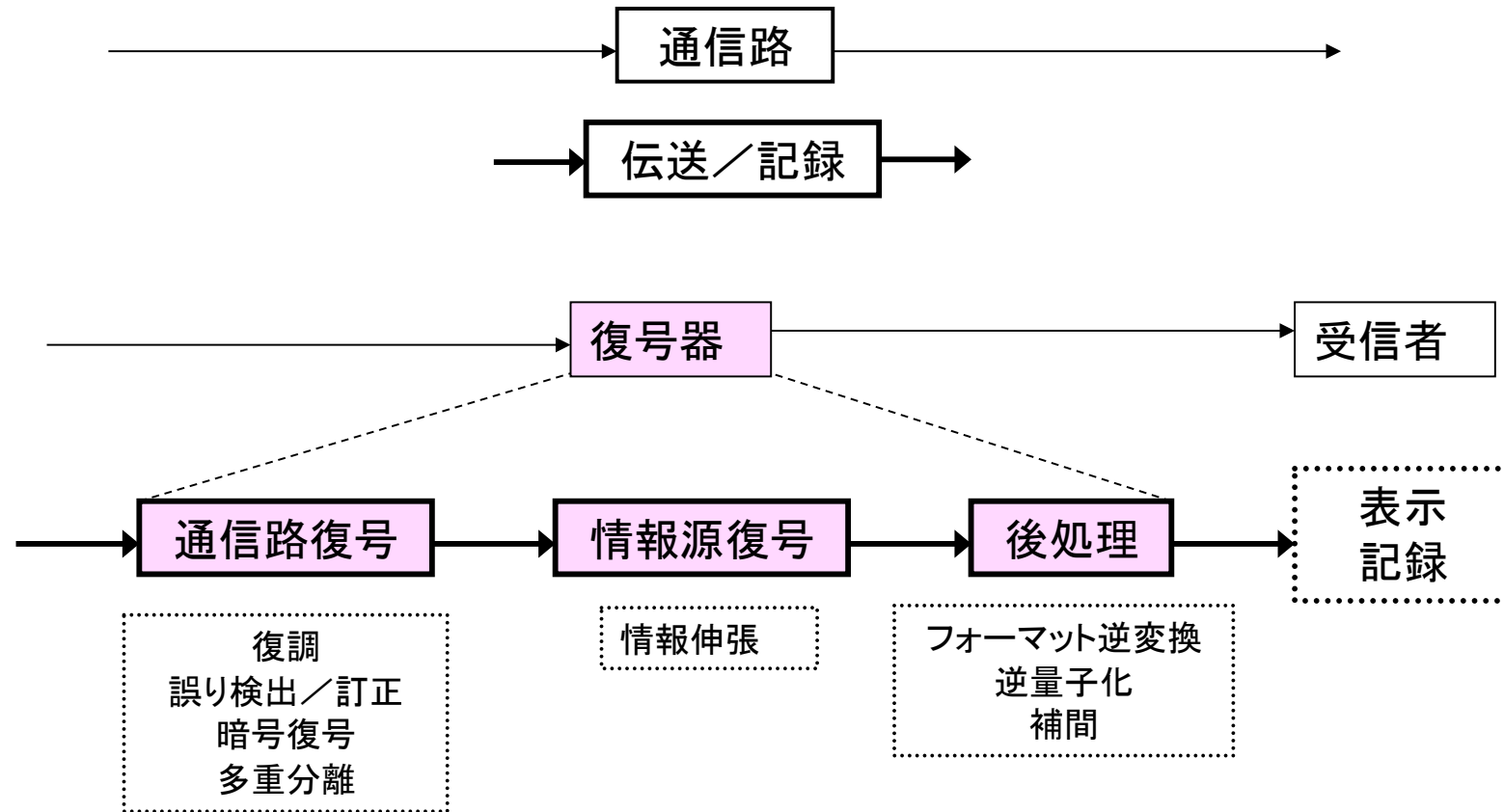
情報源 (information source) で発生した**通路** (message) が**符号器** (coder, encoder) によって**信号** (signal) に変換され、**通信路** (channel) に送出される。**復号器** (decoder) では通信路から受信した信号を通報に復元し、人間または機械に相当する**受信者** (destination) がそれを受け取る。なお、通常、通信路には**雑音** (noise) が存在して信号に影響を与えるため、符号器が送出する信号と復号器が受信する信号とが常に一致するとは限らない。



テレビカメラやスキャナから取り込んだアナログ画像信号に対して、**標本化** (sampling)、**量子化** (quantization) の前処理を施した後、各種フィルタを用いてフォーマット変換を施して「0」と「1」からなるデジタル画像信号系列を生成する。次に、デジタル画像信号系列から冗長な成分を除去し、圧縮された符号化系列に変換する。このとき情報源からの信号系列に対して予測・変換処理を施したあと、**量子化レベル** (quantization level) の縮退等を行って、信号に含まれる時間的・空間的な冗長度を除去する。さらに、シンボルの生起確率に応じてシンボルに割り当てる符号の長さを変える**可変長符号化** (VLC : variable length coding) を行って情報を圧縮する。このように、情報源からの信号系列を冗長成分の少ない符号化系列に変換することを**情報源符号化** (source coding) という。

情報圧縮された信号は、**暗号化** (enciphering) 付与、**多重化** (multiplexing) **誤り検出／訂正符号** (error detecting/correcting code) 付与、**変調** (modulation) のどの処理によって伝送路や記録媒体の特性に応じた信号に変換して送信する。これを**通信路符号化** (channel coding) という。

## デジタル画像符号化伝送システム



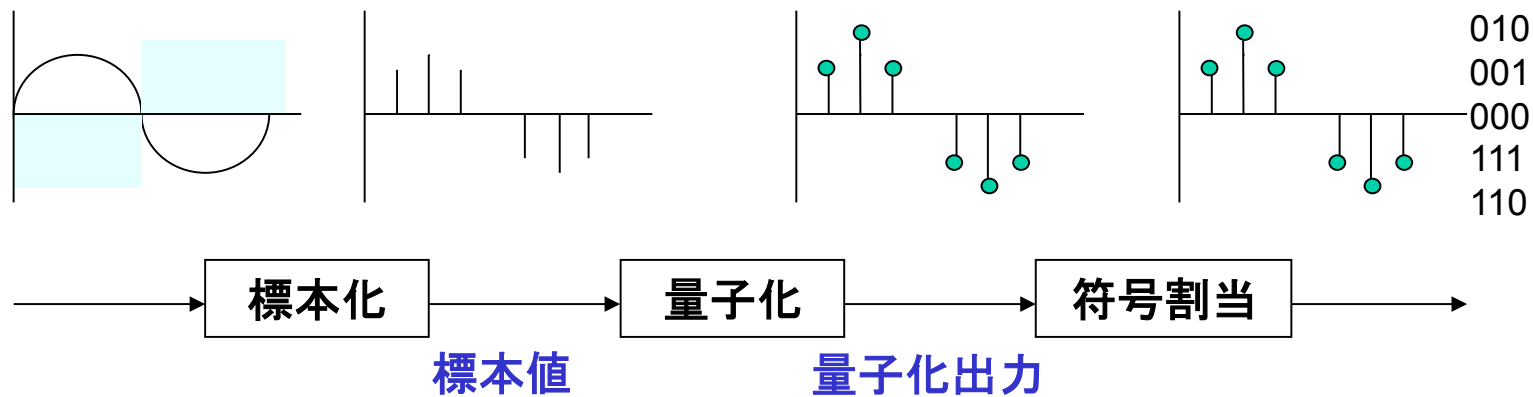
受信側では、送信側と逆の操作によって画像データが再生され、ディスプレイやハードコピーなどの出力装置から出力される。

## 4-2 アナログ信号からデジタル信号への変換

アナログ信号(連続信号)からデジタル信号(離散信号)への変換を **A/D変換**

**D変換** (Analog to Digital conversion) といい、その逆の変換を **D/A変換** (Digital to Analog conversion) という。

A/D変換の処理では、アナログ信号に対して標本化・量子化・符号割り当て(code assignment)の一連の処理行ってデジタル信号に変換する。



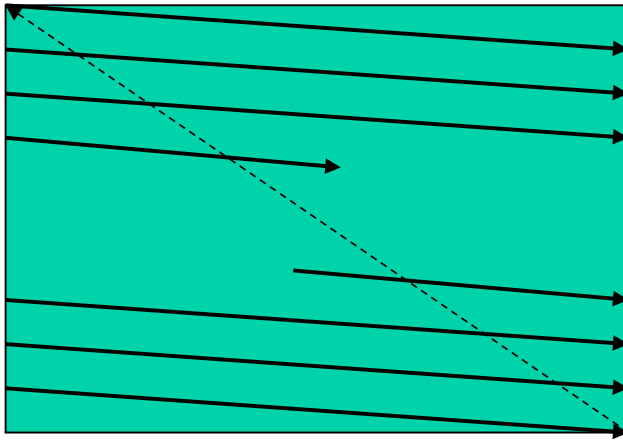
## 4-2-1 標本化

標本化は、信号を時間方向に離散化することである

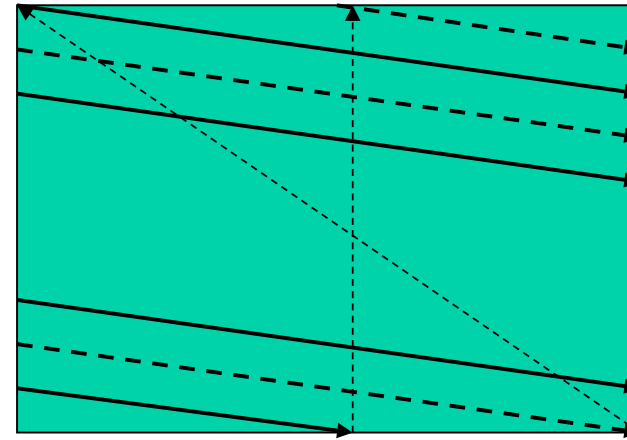
### シャノン-染谷の標本化定理

帯域が  $f_0/2$  [Hz] 以下の信号に対して、標本化周波数を  $f_0$  [Hz] 以上とすれば、歪を伴うことなく信号を復元できる

**音声** 約 3.4kHz の信号帯域を有する通常のアナログ電話音声信号をデジタル音声信号に変換する場合、標本化周波数 (sampling frequency) を 8kHz にすることが多い。また、可聴帯域 (～20kHz 程度) の音響信号の場合、44.1kHz や 48kHz などの標本化周波数が用いられる。



順次走査



飛び越し走査

## テレビ信号の走査

### 画像

画像は通常、2次元平面上に分布した情報である。動画の場合には、時間軸を追加した3次元の中の各座標点に明るさの情報(輝度情報)や色情報が存在する。このような多次元の情報をも1次元の信号系列に変換するために走査(scanning)を行う。走査方式には、画面の左上から右下に向かって走査線(scanning line)を1本1本順に走査して1枚の画面「フレーム(frame)」を構成する順次走査(progressive scanning)と、走査線を1本おきに走査し、画面の上から下に2回行って1枚のフレームを構成する飛び越し走査(interlace scanning)の2種類がある。

**NTSC** (National Television System Committee) では、飛び越し走査を行っている。走査線は 525 本 (実際画像は 480 本、残りは同期情報) で 1 フレームを構成する。フレームレートは 29.97 フレーム／秒である。フィールドレートは 59.94 フィールド／秒。

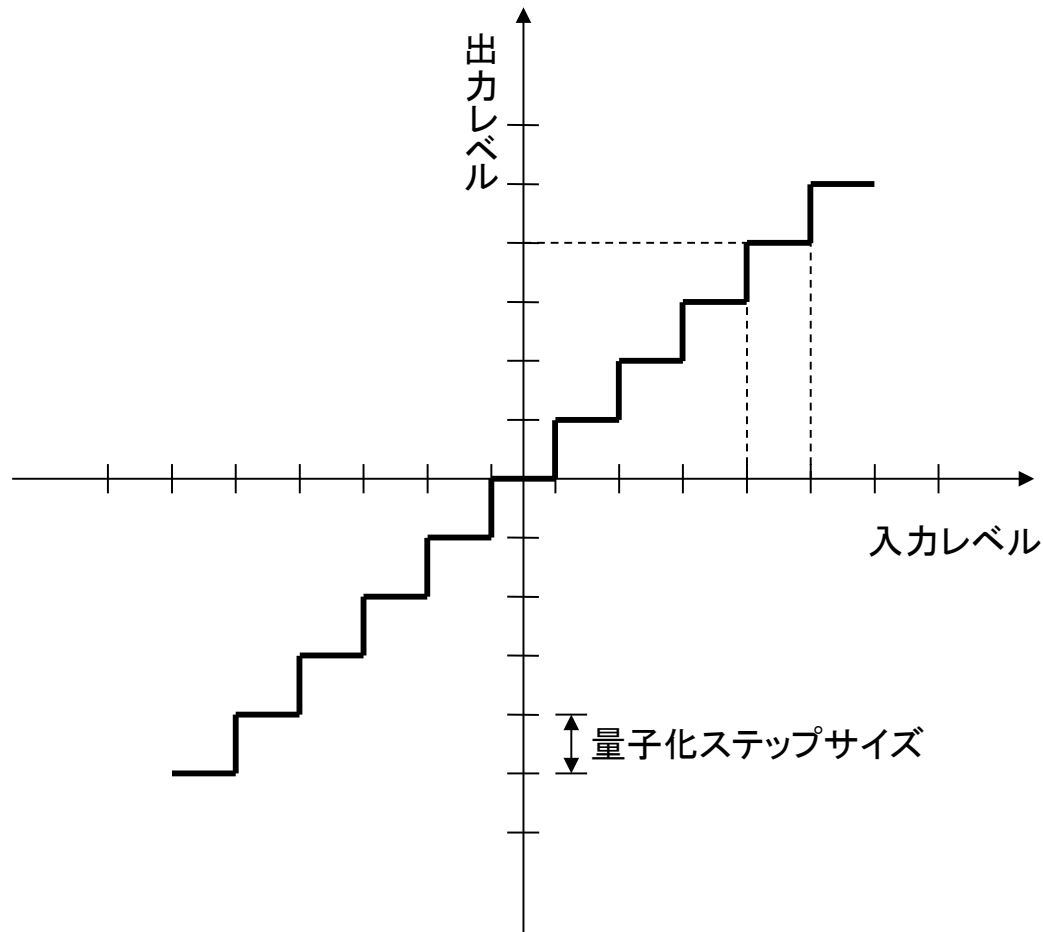
NTSC 方式では 3 原色 (RGB) を輝度信号 (luminance signal) と 2 つの色差信号 (color difference signal) に変換して伝送している。これらをデジタル信号に変換する場合、輝度信号を 13.5MHz で、色差信号を 6.75MHz で標本化している。

画像の場合、デジタル化によって得られた標本値を画素 (pel : picture element あるいは pixel : picture cell) という。有効画素数は、走査線当たり輝度信号 Y で 720 [画素／ライン]、色素信号  $C_B$ ,  $C_R$  でそれぞれ 360 [画素／ライン] である。

|  |           |
|--|-----------|
| <b>NTSC</b> (National Television System Committee) | : アメリカ、日本 |
| <b>PAL</b> (Phase Alternation by Line)             | : ヨーロッパ   |
| <b>SECAM</b> (Sequential Color and Memory)         | : アジア     |



## 4-2-2. 量子化



量子化は、信号を振幅方向に離散化することである。出力レベルを**量子化代表値** (quantization representative value) または**量子化レベル**という。また、量子化レベルの数が  $2^n$  の**量子化器** (quantizer) を  $n$  ビットの量子化器という。

入力レベルが量子化器に入力される標本値、出力レベルが量子化代表値という。量子化代表値の間隔を**量子化ステップサイズ** (quantization step size) といい、すべての量子化ステップサイズが等しいような量子化特性を持つ量子化を**線形量子化** (linear quantization) または**一様量子化** (uniform quantization) という。

## 4-2-3. 符号割当て

量子化出力値をデジタル信号として伝送するには、個々の量子化出力値に対して「0」と「1」からなる符号語(二値符号)を割り当て、これらの符号を伝送する必要がある。符号語の長さが量子化出力値(シンボル)に関わらず一定であるような符号化を**固定長符号化** (FLC : fixed length coding) という。一方、量子化出力値の生起確率が一様でない場合、伝送すべきシンボル全体の符号量を削減する上で**可変長符号化** (VLC : variable length coding) が有効である。符号化には**ハフマン符号化** (Huffman coding) や**算術符号化** (arithmetic coding) がある。

**算術符号化** : シンボルの生起確率に応じて符号語を算術演算により逐次生成する方法    **ハフ**

**マン符号化** : 生起確率が大きいシンボルに短い符号語を割り当てる方法



## 4-3-1. 情報源符号化の目的

画像の伝送には限られた帯域の伝送路を有効に使用するため情報圧縮が必要

### アナログ音声のデジタル化

標本化周波数 : 8kHz

量子化 : 8bit [= 256 レベル]、

符号割当 : 8bit 固定長符号によってデジタル化して伝送する場合

→ ビットレートは  $8\text{kHz} \times 8\text{bit} = 64\text{kb/s}$

★ 64kb/s のデジタル回線1本で伝送できる。

### NTSC の画像のデジタル化

フレームレート : 29.97 フレーム

1フレーム当たりの解像度

輝度信号 Y : 720 画素×480 ライン

色差信号 CB, CR : 360 画素×240 ライン

量子化 : 8 bit

符号割当 : 8 bit 固定長符号でデジタル化

→ ビットレートは  $((720 \times 480) + (360 \times 240) \times 2) \times 8 \times 29.97 = 165.72\text{Mbit/s}$

## 4-3-2. レートと歪の関係

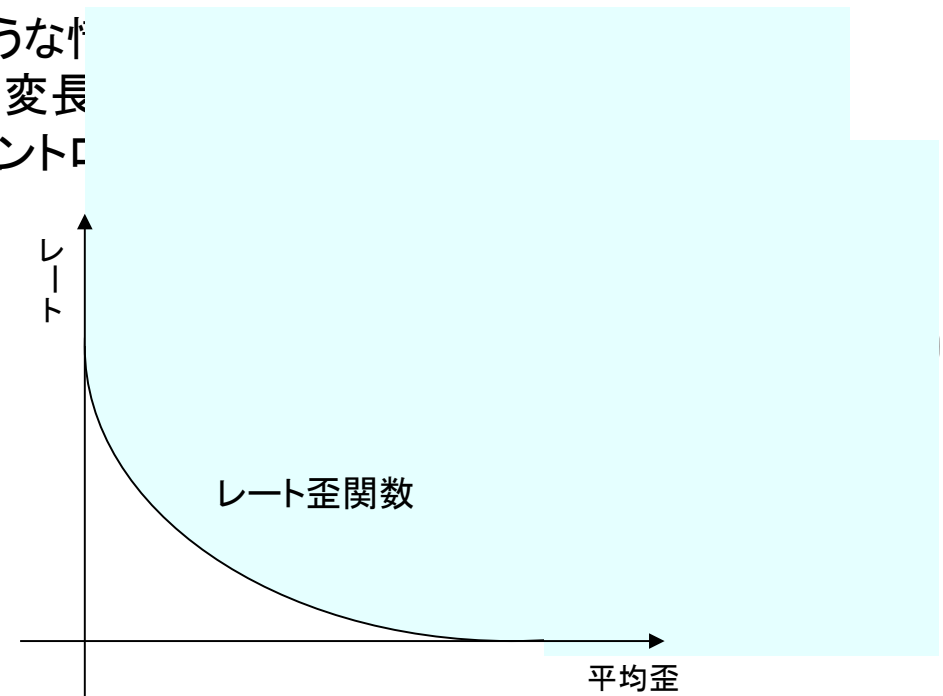
情報圧縮は、**可逆符号化** (reversible coding) と**非可逆符号化** (irreversible coding) に大別される。

**可逆符号化：**

- ★ 圧縮したデータから元の情報を完全に再現できる。  
固定長符号化と可変長符号化がある。
- ★ シンボルの生起確率が互いに独立であるような情報源 (memoryless information source) に対する可変長符号化 (entropy coding) が理論的な圧縮限界となるためエントロピー符号化 (entropy coding) とも呼ばれる

**非可逆符号化**

- ★ 圧縮したデータから再生した情報に誤差(歪)が含まれる
- ★ 圧縮率の向上が見込める  
→ 画像や音声・音響の情報圧縮で利用



## 4－3－3. 情報源符号化の要素技術

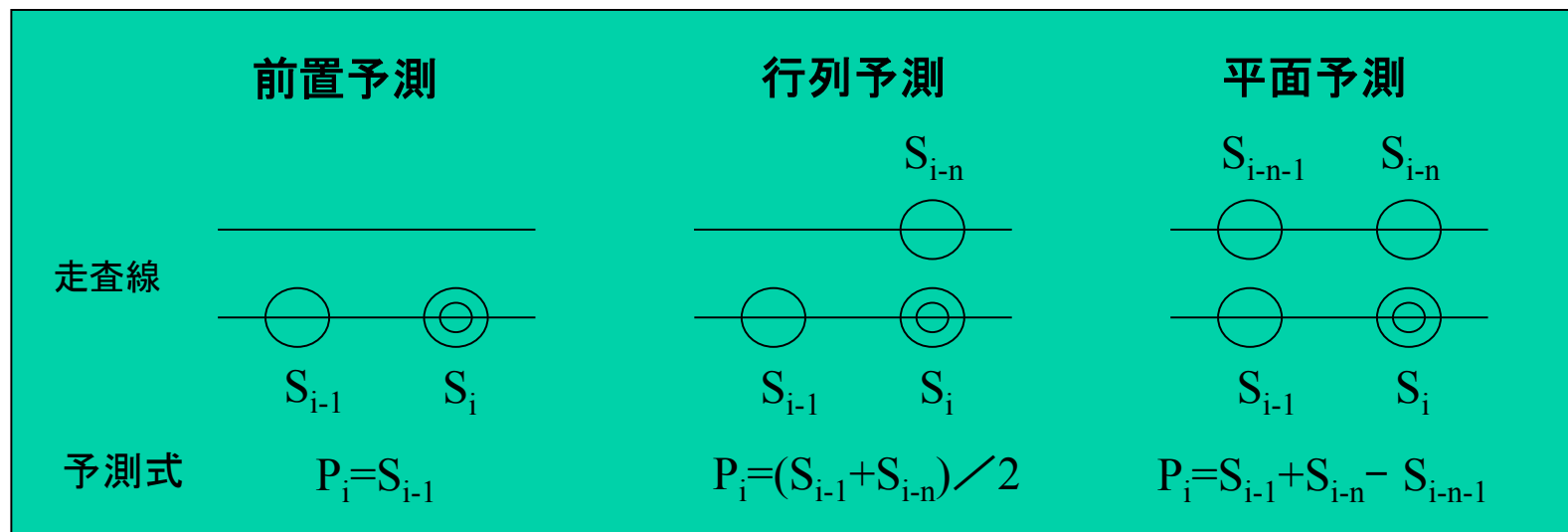
画像符号化(画像情報に対する情報源符号化)の技術

- (1) 予測符号化 (predictive coding)
- (2) 変換符号化 (transform coding)
- (3) ベクトル量子化 (VQ : vector quantization)

音声・音響符号化も同様の技術が用いられる

## (1) 予測符号化 (predictive coding)

空間的または時間的な画素間の相関性を用いて符号化すべき画素値を予測し、真の値から予測誤差を量子化して伝送する方式を**予測符号化**という。



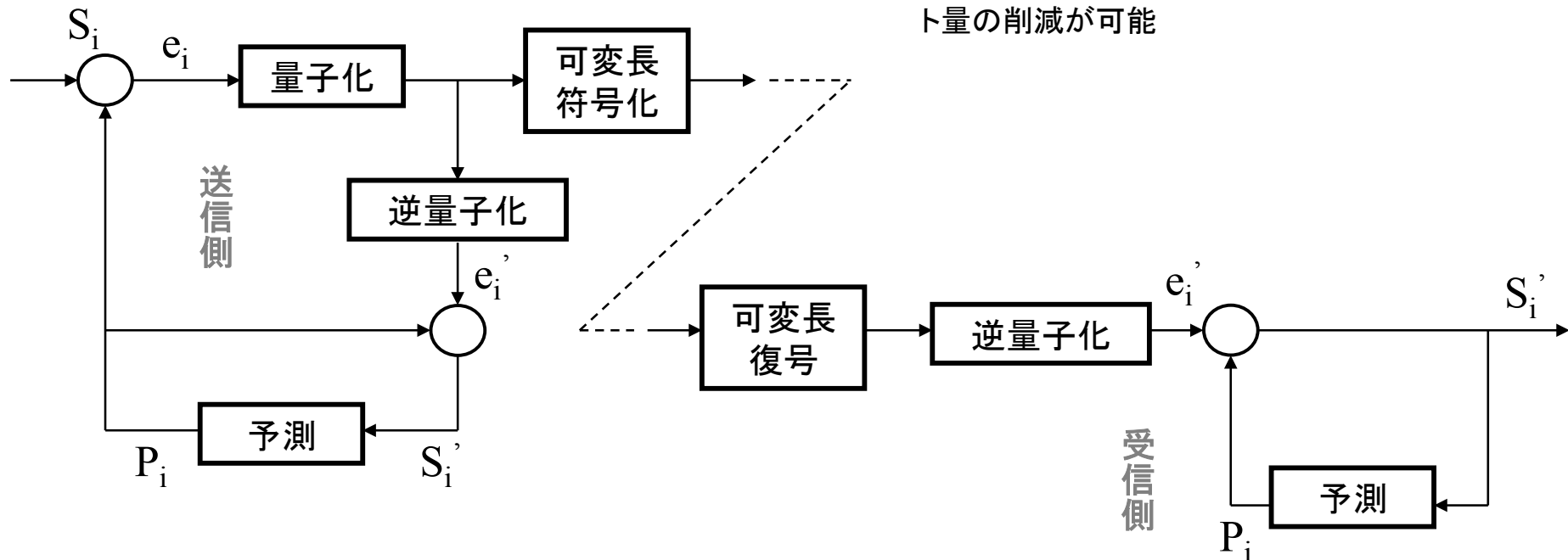
$P_i$  予測値       $n$  : 1ライン当たりの画素数

○ 予測すべき画素

◎ 予測に用いる画素(既に符号化された画素)

入力画像信号  $S_i$  を予測(近似)して予測信号  $P_i$  を生成し、入力画像信号  $S_i$  と予測信号  $P_i$  との差分によって得た予測誤差信号  $e_i$  を量子化し伝送する。予測制度が高ければ予測誤差信号  $e_i$  はゼロ近傍に集中する。従って、入力信号  $S_i$  を直接量子化するのとは比べて、量子化器の量子化レベルが少なくて済み(量子化レベルの縮退)、量子化ビット数が低減され、伝送するビット量が削減される。

可変長符号化と組み合わせると、より一層のビット量の削減が可能



逆量子化によって予測誤差信号  $e_i'$  を再生し、送信側と同じ予測によって再生された予測信号  $P_i$  と加算して復号画像信号  $S_i'$  を得る



## ① フレーム内予測 (intraframe prediction)

同一フレーム(あるいはフィールド)上の画素を用いて符号化伝送すべき画素の予測を行うこと

### ★ 前置予測

画素  $S_i$  の予測値  $P_i$  として、画素  $S_i$  と同一走査線上にある既に符号化した隣接画素  $S_{i-1}$  を用いる

### ★ 行列予測

同一走査線上および隣接する走査線上の符号化済み画素を用いて予測値  $P_i$  を生成する

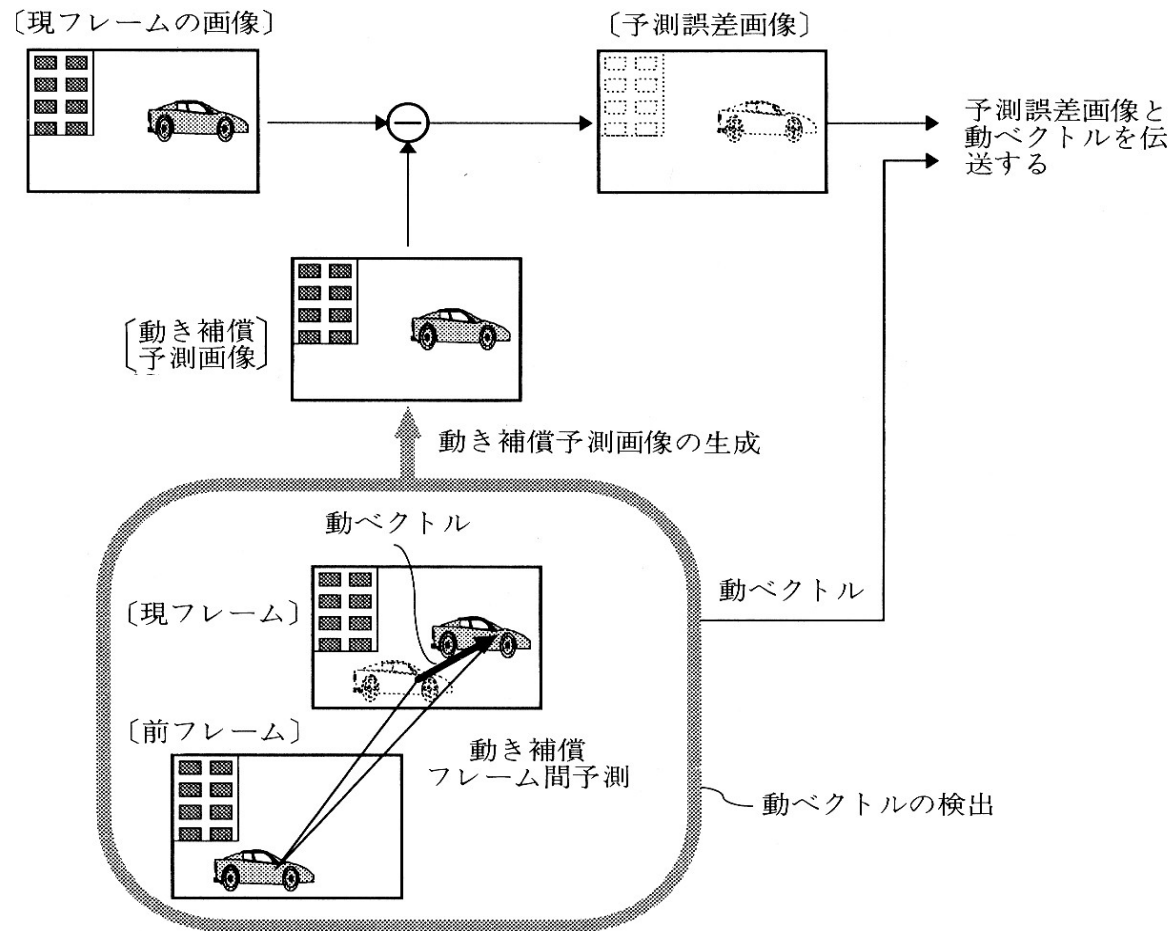
### ★ 平面予測

同一走査線上および隣接する走査線上の符号化済み画素を用いて予測値  $P_i$  を生成する

フレーム内予測では空間的に近傍にある画素同士の相関性を利用して予測値  $P_i$  を生成している

## ② フレーム間予測 (interframe prediction)

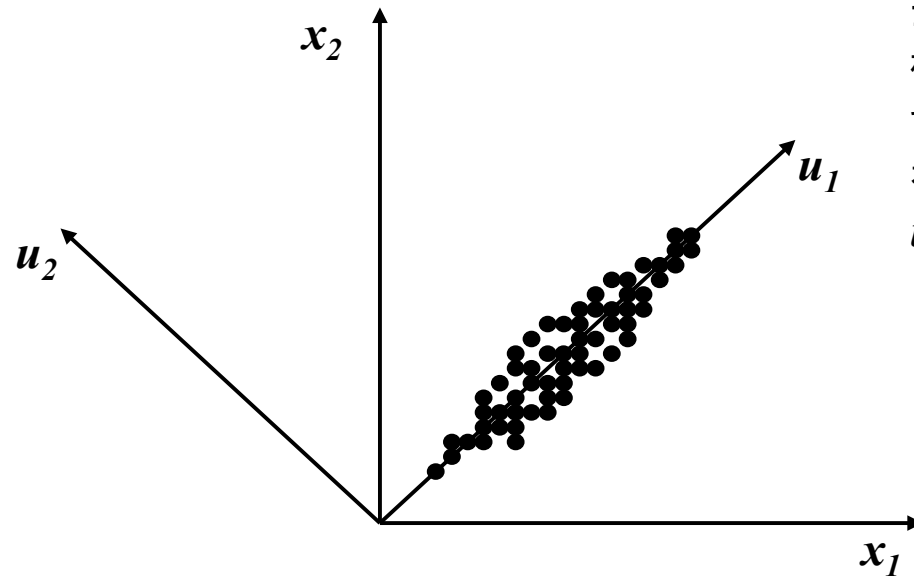
既に符号化した別のフレームの画素を用いて符号化伝送すべき画素の予測をする



フレーム間予測は時間的に近いフレーム同士の相関性を利用して予測値  $P_i$  を生成する。現フレームの画素を符号化する場合、前フレームの画素(符号化済み)との差分によって生成した予測誤差画像を量子化して伝送するのが最も簡単なフレーム間予測の方法である。

## (2) 変換符号化 (transform coding)

信号系列を変換して特定の変換出力(変換係数(transform coefficient))にパワー(変換係数の2乗)を集中させ、これらを量子化することによってビット量を削減する方式



画面上で隣接する2個の画素をまとめて符号化する場合、これらの画素の間には強い相関(隣接する画素同士はほぼ等しい値をとることが多い)があり、2次元座標( $x_1$  -  $x_2$ )に任意の隣接する2個の画素をプロットすると、左図のように直線 $x_1=x_2$ の近傍に集中して分布する。座標軸を45°回転して新しい座標軸( $u_1$  -  $u_2$ )に変換する

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

このような座標変換によってプロットは $u_1$ 軸上に集中する、そしてこの座標系で量子化すれば量子化ビット数が低減される。

画像は一般に隣接する画素間の相関性が高い。1次元あるいは2次元的に分布する画素の集合を周波数領域に変換すれば、低域側にパワーの大半が集中する傾向があり、高域側の成分に対する量子化ビット数を削減できる。

### 離散コサイン変換 (DCT : discrete cosine transform)

入力画像を水平・垂直方向とも  $N$  画素からなる  $N \times N$  画素 (通常  $8 \times 8$ ) のブロックに分割して各々の画素ブロックに対して DCT を施すとき、

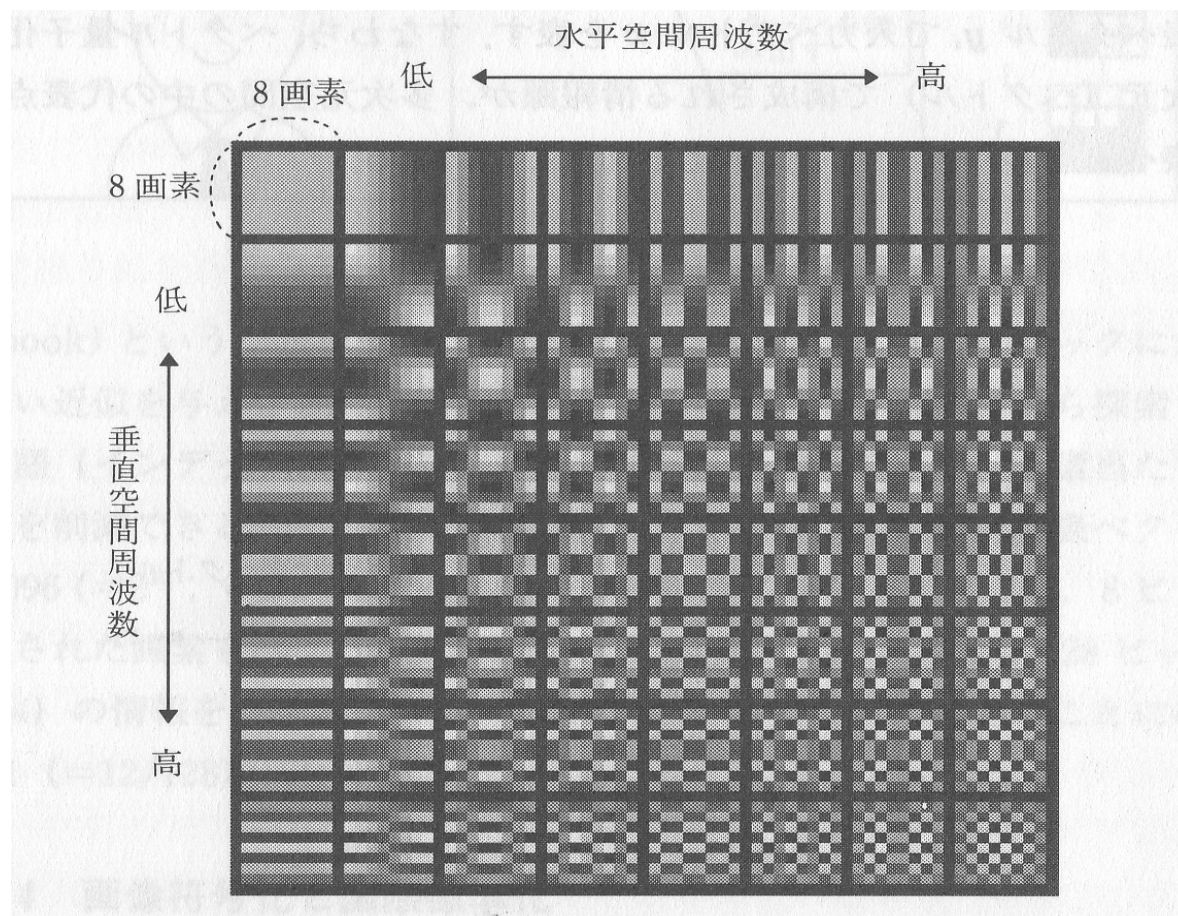
$$F(u, v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[ \frac{(2x+1)u\pi}{2N} \right] \cos \left[ \frac{(2y+1)v\pi}{2N} \right]$$
$$C(u) C(v) = \begin{cases} \frac{1}{\sqrt{2}} & (u, v = 0) \\ 1 & (u, v \neq 0) \end{cases}$$

DCT変換とは、画像データを横方向の細かさと縦方向の細かさに分解し、DCT係数というデータ値の変化の度合いを示すデータに変換することである

この変換により得られる  $F(u, v)$  をDCT係数と言い、上式のような変換を**2次元DCT**という

## 逆離散コサイン変換

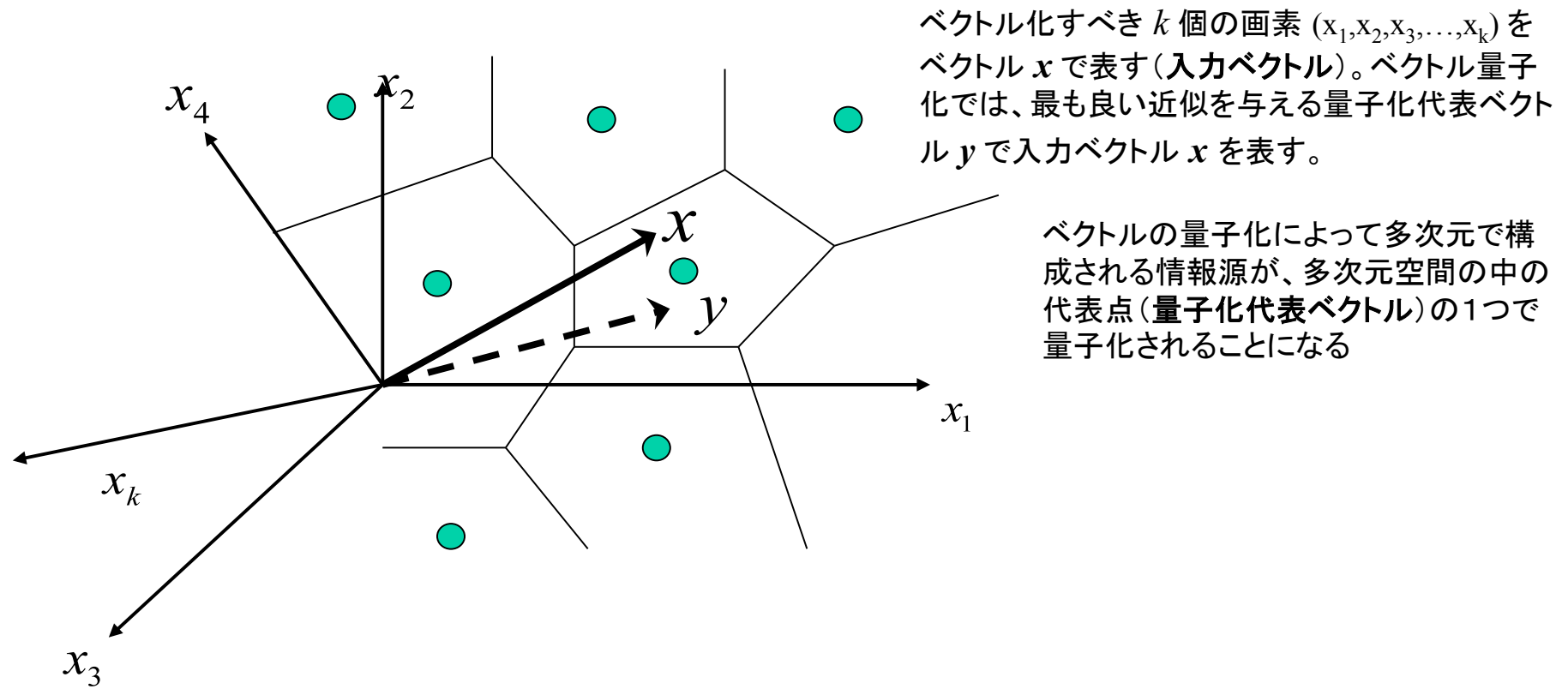
$$f(x, y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$



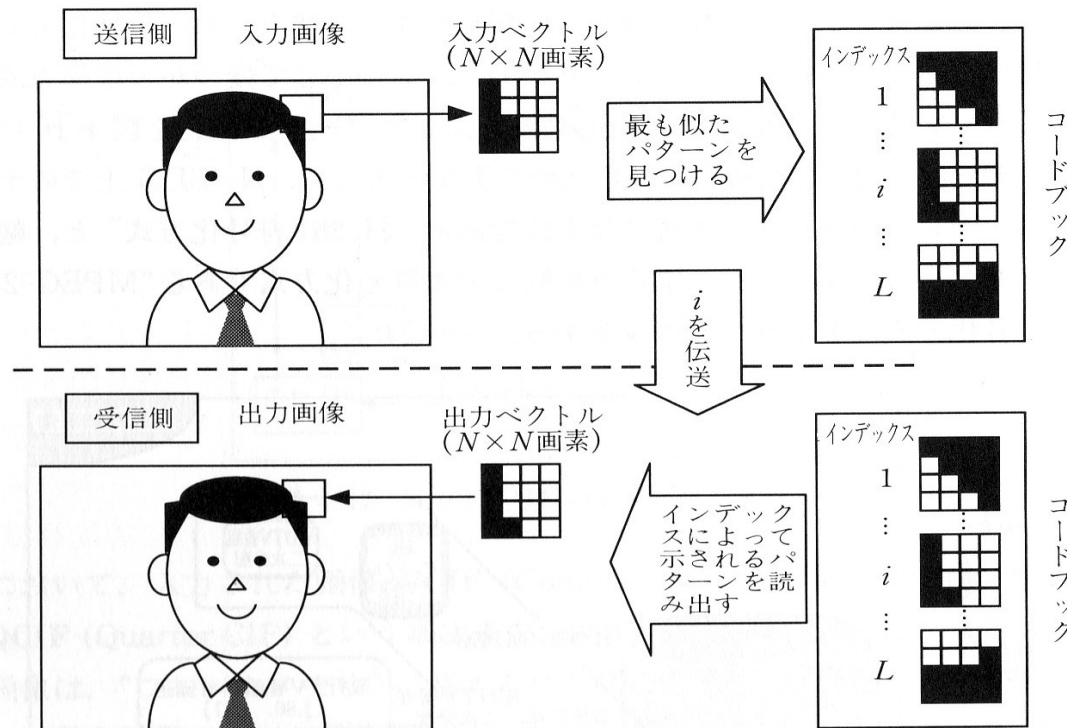
### (3) ベクトル量子化 (VQ : vector quantization)

複数個の標本をまとめて量子化することによってビット量を削減する方式をベクトルの量子化という。

「スカラー量子化 (scalar quantization) : 1次元のベクトルの量子化」



$N \times N$ 画素のブロックをベクトル量子化する場合、送信側、受信側で予め用意しておく量子化代表ベクトルの集合をコードブック(codebook)といい、送信側では、符号化すべき $N \times N$ 画素ブロックに対して最も良い近似を与える量子化代表ベクトルをコードブックの中から探索し、その符号語(インデックス)を受信側へ伝送することで、画素当たりのビット量を削減できる。



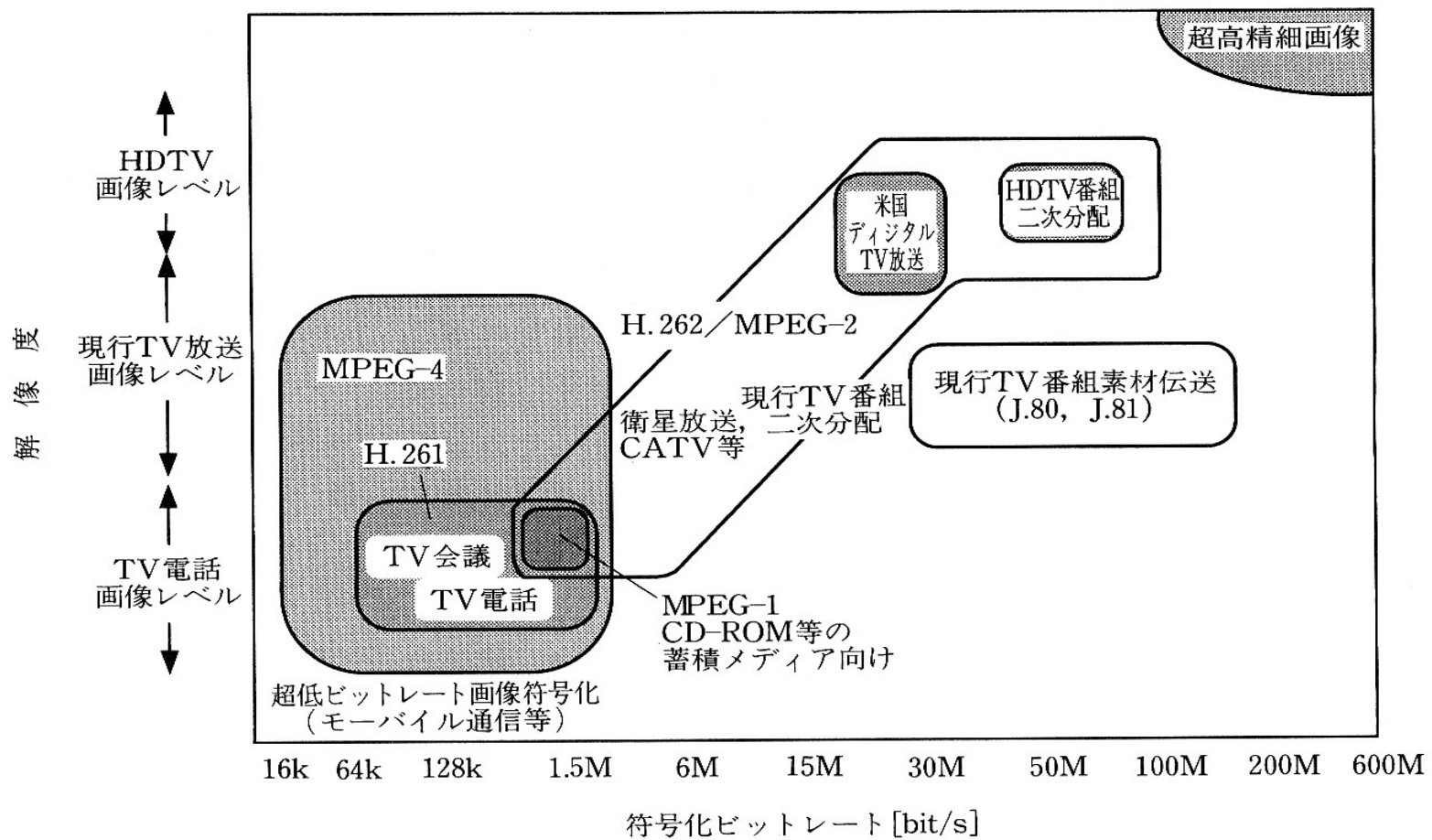
$N=4$ 、コードブック中の量子化代表ベクトルの数を4096 ( $2^{12}$ : 12ビット長のインデックス)として、8ビットで量子化された画素をベクトル量子化すると、画素当ブロックたり128ビット ( $8 \times 4 \times 4$ ) の情報を、画素ブロック当たり12ビットで伝送出来る事になり、約  $1/11 (= 12/128)$  に圧縮できる。



## 4-3-4. 画像符号化と国際標準化

デジタル画像の符号化に関する要素技術は、画像信号がもつ種々の相関性を利用して冗長な成分を削減するもので、高能率符号化技術(high-efficiency coding)と呼ばれる。画像の高能率符号化技術の実用化においては、画像コンテンツの相互流通性・相互操作性の観点から、符号化方式を国際標準化することが重要である。アプリケーションに応じて幾つかの標準方式が定められている。主な標準方式は、

- (1) H.261符号化方式: N-ISDN上でのテレビ電話・テレビ会議用ビデオ符号化方式
- (2) MPEG-2ビデオ符号化方式: 汎用的ビデオ符号化方式



## (1) H. 261 ビデオ符号化方式

H.261符号化方式は、国際電信電話諮問委員会(CCITT: International Telephone and Telegraph Consultative Committee)において1990年に勧告化され、N-ISDN(64k～2Mbit/s)上でのテレビ電話・テレビ会議用として使用される動画像符号化方式である。

**CIF/QCIF変換**においては、入力された画像がCIF (Common Intermediate Format) または QCIF (Quarter CIF) という順次走査の画像に変換される。「CIFでは輝度信号：352画素×288ライン、色差信号：176画素×144ラインが符号化対象領域で、フレームレートは29.97フレーム／秒。QCIFはCIFの1／4 (水平・垂直それぞれ1／2) でフレームレートは同じ」。CIF または QCIF 画像は「フレームメモリ」に蓄えられた符号化 (および複合化) 済みの信号を用いて動き補償フレーム間予測が行われ、予測誤差信号がDCTによってDCT係数に変換された後、量子化される。量子化された係数は符号化情報として出力されると同時に量子化代表値に復元され、IDCTによって予測誤差信号が再生される。この予測誤差信号が先の予測信号と加算され、画像信号が復号される。

**可変長符号化**においては、変換係数の量子化結果や動ベクトルなどの符号化データ、各種パラメータ、付加情報などを可変長符号化し多重化する。

**バッファ**においては、符号化ビット列を一時的に蓄積して一定のビットレートで伝送路に送出する。

**符号化制御**においては、バッファ蓄積量等に基づいて、フレーム間／フレーム内モードや量子化ステップサイズなどの切り換えや、予測の際の広域の歪を低減するために設けた低域通過型フィルタのオン／オフ制御を行う

**誤り訂正符号化**においては、伝送誤りの影響を抑えるために誤り訂正符号化を行う。

復号器ではほぼその逆

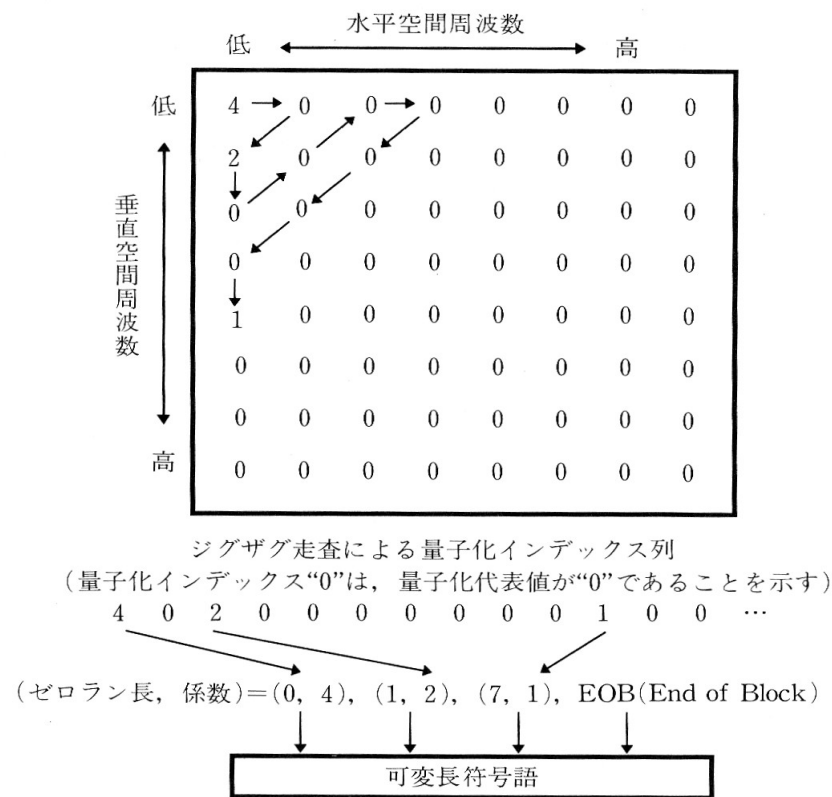


図 4.18 ジグザグ走査と 2 次元 VLC

処理としては、ジグザグ走査と2次元VLCがある。

画像信号にDCTを施した場合、低域側にパワーが片寄り、高域側ほどゼロに近くなる。そこで、DCT係数の量子化インデックスをブロック内部の低域側から高域側へジグザグに走査する。このような走査をジグザグ走査という。ジグザグ走査を行うと、走査の始めではゼロ以外に量子化されたインデックスが連続して現れ、走査の終わりにいくに従ってゼロに量子化されたインデックスが連続する。このとき、ブロック内部で未走査のまま残っている量子化インデックスがすべてゼロであれば、EOBのコードを付けて以下のゼロインデックスを無視する。これにより、伝送すべきビット量を削減できる。走査途中では、ゼロに


量子化されたインデックスが連続する数(ゼロランの数)と、それに続く非ゼロなるインデックスの値とを組み合わせた事象に対して可変長符号語を割り当てる。このような可変長符号化の方法を2次元VLCという。

## (2) MPEG-2ビデオ符号化方式

MPEG(Moving Picture Experts Group)の標準方式はビデオ、オーディオ、システム等のパートに分かれており、それぞれビデオ符号化方式、オーディオ符号化方式、ビデオやオーディオ等のメディアの多重化および同期に関する方式を規定している。

MPEG-2ビデオ符号化方式では、各アプリケーションに対応した装置が複雑になりすぎないように、また、符号化情報の交換を容易にするように、プロファイルのレベルによる段階付けをしている。

|  |                    |                      |                     |                       |                      |                     |
|--|--------------------|----------------------|---------------------|-----------------------|----------------------|---------------------|
| ハイレベル<br>~1920画素/line<br>~1152line     |                    | MP@HL<br>~80Mbit/s   |                     |                       | HP@HL<br>~100Mbit/s  |                     |
| ハイ1440レベル<br>~1440画素/line<br>~1152line |                    | MP@H-14<br>~60Mbit/s |                     | Spt@H-14<br>~60Mbit/s | HP@H-14<br>~80Mbit/s |                     |
| メインレベル<br>~720画素/line<br>~576line      | SP@ML<br>~15Mbit/s | MP@ML<br>~15Mbit/s   | SNR@ML<br>~15Mbit/s |                       | HP@ML<br>~20Mbit/s   | 422@ML<br>~50Mbit/s |
| ローレベル<br>~352画素/line<br>~288line       |                    | MP@LL<br>~4Mbit/s    | SNR@LL<br>~4Mbit/s  |                       |                      |                     |

 は未使用

|                           |                          |                                    |                                   |                               |                                  |
|---------------------------|--------------------------|------------------------------------|-----------------------------------|-------------------------------|----------------------------------|
| シンプル<br>プロファイル<br>(4:2:0) | メイン<br>プロファイル<br>(4:2:0) | SNR<br>スケーラブル<br>プロファイル<br>(4:2:0) | 空間<br>スケーラブル<br>プロファイル<br>(4:2:0) | ハイ<br>プロファイル<br>(4:2:2/4:2:0) | 4 2 2<br>プロファイル<br>(4:2:2/4:2:0) |
|---------------------------|--------------------------|------------------------------------|-----------------------------------|-------------------------------|----------------------------------|