信号理論

- No.9 A/D変換と量子化 -

渡辺　裕

Signal Theory

- No.9 A/D Conversion and Quantization -

Hiroshi Watanabe

A/D 変換

- アナログ信号からディジタル信号へ
- プレフィルタ処理：高周波成分の除去
- 標本化：時間軸における離散化
- 量子化：レベル値の有限n進数への変換

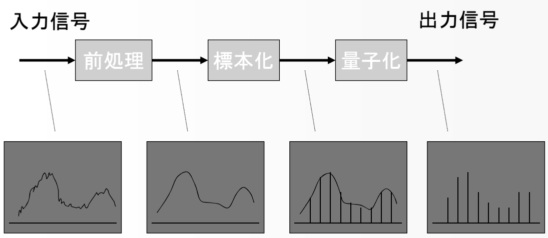
アナログ信号 → プレフィルタ → 標本化 → 量子化 → ディジタル信号

A/D Conversion

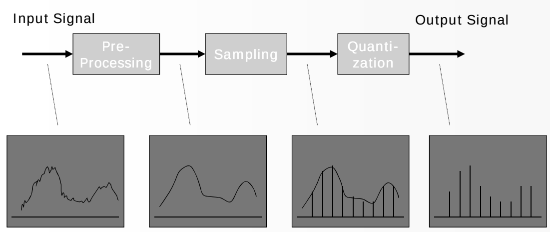
- Analog Signal to Digital Signal
- Pre-filtering: Cut High Frequency Component
- Sampling: Discrete in Time Domain
- Quantization: Convert Value to Finite Digits

Analog Signal → Pre-Filtering → Sampling → Quanti-zation → Digital Signal

処理ダイヤグラム

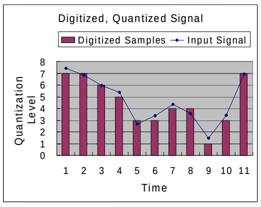


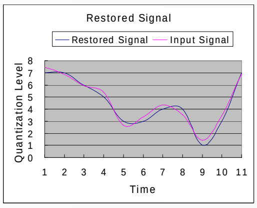
Processing Diagram



PCM

- PAMデータへのユニーク符号の割り当て

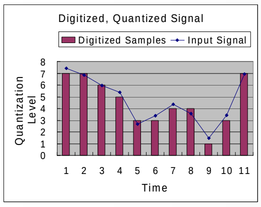


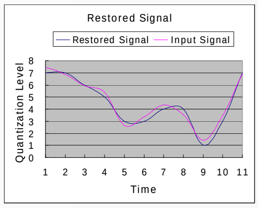


離散化・量子化信号　　　再生信号

PCM

- Unique Code Assignment to PAM Data





Digitized and quantized signal　　Restored signal

2進数と10進数

- 2進数表現

$$b_n b_{n-1}\cdots b_1 b_0 . b_{-1} b_{-2}\cdots b_{-m}\Big|_2 = \left. \begin{array}{l} b_n 2^n + b_{n-1}2^{n-1}+\cdots+b_1 2^1 + b_0 2^0 \\ + b_{-1}2^{-1}+\cdots+b_{-m}2^{-m}\end{array}\right|_{10}$$

$$= \left.\sum_{i=-m}^{n} b_i 2^i\right|_{10} , \quad b_i = \begin{cases}0\\1\end{cases}$$

- 例

$$10110 = 2^4 + 2^2 + 2^1 = 16 + 4 + 2 = 22$$

Binary and Decimal Numbers

- Binary Representation

$$b_n b_{n-1}\cdots b_1 b_0 . b_{-1} b_{-2}\cdots b_{-m}\Big|_2 = \left. \begin{array}{l} b_n 2^n + b_{n-1}2^{n-1}+\cdots+b_1 2^1 + b_0 2^0 \\ + b_{-1}2^{-1}+\cdots+b_{-m}2^{-m}\end{array}\right|_{10}$$

$$= \left.\sum_{i=-m}^{n} b_i 2^i\right|_{10} , \quad b_i = \begin{cases}0\\1\end{cases}$$

- Example

$$10110 = 2^4 + 2^2 + 2^1 = 16 + 4 + 2 = 22$$

2進符号

- レベル数とビット数

Level数	Bit	Code
2	1	0, 1
4	2	00, 01, 10, 11
8	3	000, 001, 010, 011, 100, 101, 110, 111
...		
2^n	n	00...0, ...

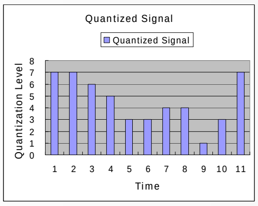
Binary Code

- Level and Bit Numbers

Level#s	Bit	Code
2	1	0, 1
4	2	00, 01, 10, 11
8	3	000, 001, 010, 011, 100, 101, 110, 111
...		
2^n	n	00...0, ...

PCM符号の割り当て

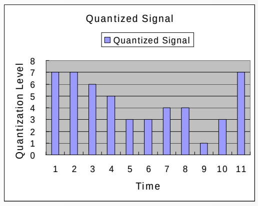
Level	Codeword
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111



時刻:	1	2	3	4	5	6	7	8	9	10	11
レベル:	7	7	6	5	3	3	4	4	1	3	7
符号:	111	111	110	101	011	011	100	100	001	011	111 / 33 bit

Assignment of PCM Code

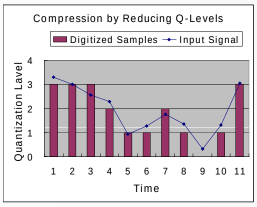
Level	Codeword
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111



Time:	1	2	3	4	5	6	7	8	9	10	11
Level:	7	7	6	5	3	3	4	4	1	3	7
Code:	111	111	110	101	011	011	100	100	001	011	111 / 33 bit

基本的なデータ圧縮 – 量子化レベル数の削減

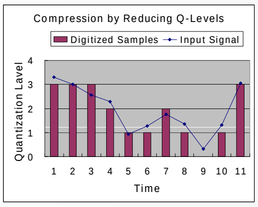
Level	Codeword
0	00
1	01
2	10
3	11



時刻:	1	2	3	4	5	6	7	8	9	10	11
レベル:	3	3	3	2	1	1	2	1	0	1	3
符号:	11	11	11	10	01	01	10	10	00	10	11 / 22bit

Basic Data Compression – Reducing Number of Quantization Levels

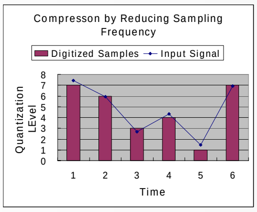
Level	Codeword
0	00
1	01
2	10
3	11



Time:	1	2	3	4	5	6	7	8	9	10	11
Level:	3	3	3	2	1	1	2	1	0	1	3
Code:	11	11	11	10	01	01	10	10	00	10	11 / 22bit

基本的なデータ圧縮 – 標本点数の削減

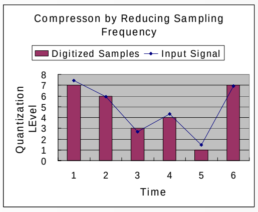
Level	Codeword
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111



時刻:	1	2	3	4	5	6
レベル:	7	6	3	4	1	7
符号:	111	110	011	100	001	111 / 18 bit

Basic Data Compression – Reducing Number of Samples

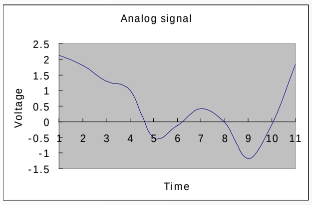
Level	Codeword
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111



Time:	1	2	3	4	5	6
Level:	7	6	3	4	1	7
Code:	111	110	011	100	001	111 / 18 bit

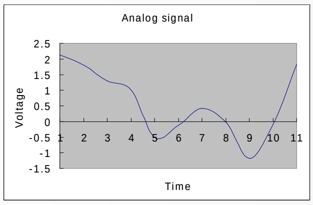
定量的解析

時刻	アナログ信号値
0	2.12
1	1.80
2	1.30
3	1.00
4	-0.50
5	-0.12
6	0.42
7	-0.02
8	-1.17
9	-0.08
10	1.84



Quantitative Analysis

Time	Analogue signal Value
0	2.12
1	1.80
2	1.30
3	1.00
4	-0.50
5	-0.12
6	0.42
7	-0.02
8	-1.17
9	-0.08
10	1.84



定量的解析 (2)

- 量子化ステップ d の算出

$$d = 2V/L$$

V: 入力電圧の最大値
L: 量子化レベル数

$L = 2^n$ 　 n : 符号のビット数

- 例

V=2.2, n=4 のとき L=16 であるから

$$d = 0.275$$

Quantitative Analysis (2)

- Calculation of Quantization step d

$$d = 2V/L$$

V: Maximum value of input
L: number of quantization levels

$L = 2^n$ 　 n : number of bits for code

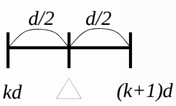
- Example

When V=2.2, n=4, L *equals to* 16, thus,

$$d = 0.275$$

定量的解析 (3)

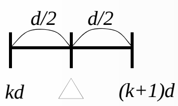
- 量子化誤差電力: σ_q^2



$$\sigma_q^2 = \frac{1}{d}\int_{-d/2}^{d/2} x^2 dx = \frac{1}{d}\left[\frac{1}{3}x^3\right]_{-d/2}^{d/2} = d^2/12$$

Quantitative Analysis (3)

- Quantization Error Power: σ_q^2



$$\sigma_q^2 = \frac{1}{d}\int_{-d/2}^{d/2} x^2 dx = \frac{1}{d}\left[\frac{1}{3}x^3\right]_{-d/2}^{d/2} = d^2/12$$

定量的解析 (4)

- 量子化によるSNR(雑音対信号比)

$$SNR(dB) = 20\log_{10}\frac{\sigma_x}{\sigma_q} = 10\log_{10}\frac{\sigma_x^2}{\sigma_q^2}$$

- 入力信号の分散: σ_x^2 (画像処理では最大値)

$$\sigma_x^2 = 4V^2$$

Quantitative Analysis (4)

- SNR (Signal to Noise Ratio) caused by Quantization

$$SNR(dB) = 20\log_{10}\frac{\sigma_x}{\sigma_q} = 10\log_{10}\frac{\sigma_x^2}{\sigma_q^2}$$

- Variance of Input signal: σ_x^2 (maximum value for image processing)

$$\sigma_x^2 = 4V^2$$

定量的解析 (5)

- ***n*** ビット量子化時のSNRの算出

$$\begin{aligned} SNR(dB) &= 10\log_{10}\frac{4V^2}{d^2/12} \\ &= 10\log_{10}\frac{4V^2}{4V^2/12L^2} \\ &= 10\log_{10}12L^2 \\ &= 10(\log_{10}3 + 2\log_{10}2) + 20\log_{10}2^n \\ &= 10.79 + 6.02n \end{aligned}$$

$$\log_{10}2 \cong 0.301 \qquad \log_{10}3 \cong 0.477$$

*1bit 6dB*の法則

Quantitative Analysis (5)

- Calculation of SNR at *n* bit quantization

$$\begin{aligned} SNR(dB) &= 10\log_{10}\frac{4V^2}{d^2/12} \\ &= 10\log_{10}\frac{4V^2}{4V^2/12L^2} \\ &= 10\log_{10}12L^2 \\ &= 10(\log_{10}3 + 2\log_{10}2) + 20\log_{10}2^n \\ &= 10.79 + 6.02n \end{aligned}$$

$$\log_{10}2 \cong 0.301 \qquad \log_{10}3 \cong 0.477$$

Rule: *1bit 6dB*

量子化

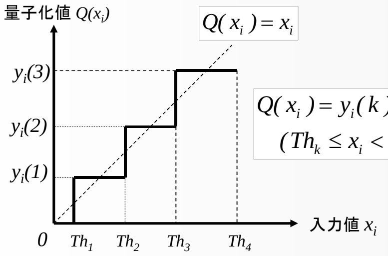
- 量子化の役割
  - 状態数の削減によるデータ圧縮
  - 量子化歪:量子化することにより生じる誤差
- 量子化の種類
  - スカラー量子化、ベクトル量子化
  - 線形量子化、非線形量子化
- 量子化の最適化
  - Max量子化器
- 画像符号化における量子化
  - JPEG, H.261, MPEG-1, MPEG-2

Quantization

- Role of Quantization
  - Data compression by reducing the number of status
  - Quantization Error: Error caused by quantization
- Types of Quantization
  - Scalar Quantization, Vector Quantization
  - Linear Quantization, Non-linear Quantization
- Optimization of Quantization
  - Max Quantizer
- Quantizer used at Image Coding
  - JPEG, H.261, MPEG-1, MPEG-2

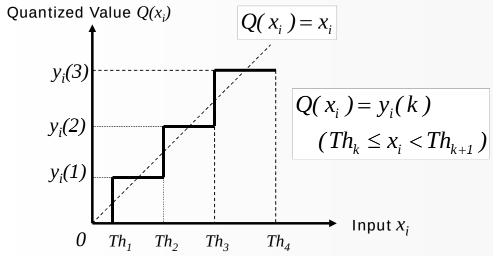
量子化とは?

- 離散レベル化



What's Quantization?

- Discrete Level



量子化歪

- スカラの場合の量子化歪と距離尺度 (L^p ノルム)

$$d(x,y) = \|x-y\|_p$$
$$\|x-y\|_p = \left(|x-y|^p\right)^{1/p}$$

- ベクトルの場合の量子化歪とノルム

$$d(\mathbf{X},\mathbf{Y}) = \|\mathbf{X}-\mathbf{Y}\|_p$$
$$\|\mathbf{X}-\mathbf{Y}\|_p = \left(|x_1-y_1|^p + |x_2-y_2|^p + \cdots + |x_n-y_n|^p\right)^{1/p}$$

Quantization Distortion

- Quantization Error and Distortion Measure for Scalar's Case (L^p-norm)

$$d(x,y) = \|x-y\|_p$$
$$\|x-y\|_p = \left(|x-y|^p\right)^{1/p}$$

- Quantization Error and Distortion Measure for Vector's Case (L^p-norm)

$$d(\mathbf{X},\mathbf{Y}) = \|\mathbf{X}-\mathbf{Y}\|_p$$
$$\|\mathbf{X}-\mathbf{Y}\|_p = \left(|x_1-y_1|^p + |x_2-y_2|^p + \cdots + |x_n-y_n|^p\right)^{1/p}$$

ノルム

- 代表的なノルム

$$\|x\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p} \quad (\infty > p \ge 1)$$
$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$
$$\|x\|_2 = \left(|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2\right)^{1/2} \quad (\text{Euclidean norm})$$
$$\|x\|_\infty \equiv \max_i |x_i|$$

Norm

- Typical Norms

$$\|x\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p} \quad (\infty > p \ge 1)$$
$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$
$$\|x\|_2 = \left(|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2\right)^{1/2} \quad (\text{Euclidean norm})$$
$$\|x\|_\infty \equiv \max_i |x_i|$$

スカラ量子化とベクトル量子化

- スカラ量子化
  - N個の要素を持つ入力信号（N次元ベクトル）$\boldsymbol{X}$ に対して、N個の量子化値を要素として持つ出力信号は、個々の要素の量子化値を並べたもの

$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$
$$Q_s(\boldsymbol{X}) = [Q_s(x_1) \quad Q_s(x_2) \quad \cdots \quad Q_s(x_N)]^t$$

$$Q_s(x_i) = (y_i(k_i) \mid k_i = \arg\min_j d(x_i, y_i(j)) \quad (j = 1,2,\ldots,r_i)$$
$$d(x,y) = \|x-y\|_p$$

信号理論 / Signal Theory 37

Scalar and Vector Quantization

- Scalar Quantization
  - For input signals (N-dimensional vector) having N elements, output signals have N elements which are quantized individually

$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$
$$Q_s(\boldsymbol{X}) = [Q_s(x_1) \quad Q_s(x_2) \quad \cdots \quad Q_s(x_N)]^t$$

$$Q_s(x_i) = (y_i(k_i) \mid k_i = \arg\min_j d(x_i, y_i(j)) \quad (j = 1,2,\ldots,r_i)$$
$$d(x,y) = \|x-y\|_p$$

信号理論 / Signal Theory 38

スカラ量子化とベクトル量子化 (2)

- ベクトル量子化
  - N個の要素を持つ入力信号（N次元ベクトル）X に対して、準備した代表ベクトル群(コードブック) $Y(i)\ (i=1,\ldots,M)$の中から最も距離の近いものを選び出す

$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$
$$Q_V(\boldsymbol{X}) = [y_1(k) \quad y_2(k) \quad \cdots \quad y_N(k)]^t$$

$$Q_V(\boldsymbol{X}) = (\mathbf{Y}(k) \mid k = \arg\min_i d(\boldsymbol{X}, \mathbf{Y}(i)) \quad (i = 1,2,\ldots,M)$$
$$\mathbf{Y}(i) = [y_1(i) \quad y_2(i) \quad \cdots \quad y_N(i)]^t$$
$$d(\boldsymbol{X}, \mathbf{Y}(i)) = \|\boldsymbol{X} - \mathbf{Y}(i)\|_p$$

信号理論 / Signal Theory 39

Scalar and Vector Quantization (2)

- Vector Quantization
  - For input signals (N-dimensional vector) having N elements, a set of output signals giving the minimum distortion is selected from codevectors stored in a codebook $Y(i)\ (i=1,\ldots,M)$

$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$
$$Q_V(\boldsymbol{X}) = [y_1(k) \quad y_2(k) \quad \cdots \quad y_N(k)]^t$$

$$Q_V(\boldsymbol{X}) = (\mathbf{Y}(k) \mid k = \arg\min_i d(\boldsymbol{X}, \mathbf{Y}(i)) \quad (i = 1,2,\ldots,M)$$
$$\mathbf{Y}(i) = [y_1(i) \quad y_2(i) \quad \cdots \quad y_N(i)]^t$$
$$d(\boldsymbol{X}, \mathbf{Y}(i)) = \|\boldsymbol{X} - \mathbf{Y}(i)\|_p$$

信号理論 / Signal Theory 40

スカラ量子化とベクトル量子化 (3)

- ベクトル量子化の概念図

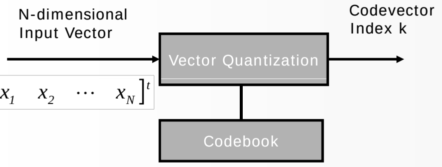
$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$

$$\mathbf{Y}(i) = [y_1(i) \quad y_2(i) \quad \cdots \quad y_N(i)]^t$$
$$(i = 1,2,\cdots,M)$$

信号理論 / Signal Theory 41

Scalar and Vector Quantization (3)

- Block Diagram of Vector Quantization



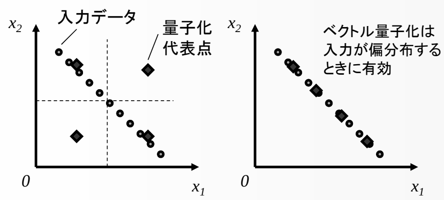
$$\boldsymbol{X} = [x_1 \quad x_2 \quad \cdots \quad x_N]^t$$

$$\mathbf{Y}(i) = [y_1(i) \quad y_2(i) \quad \cdots \quad y_N(i)]^t$$
$$(i = 1,2,\cdots,M)$$

信号理論 / Signal Theory 42

スカラ量子化とベクトル量子化 (4)

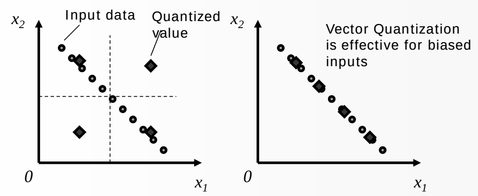
- 2次元入力のスカラ量子化とベクトル量子化



スカラ量子化　　ベクトル量子化

Scalar and Vector Quantization (4)

- Scalar and Vector Quantization for 2-D input



Scalar Quantization　　Vector Quantization

ベクトル量子化

- コードブックの作成... LBGアルゴリズム (Linde, Buzo, Gray)
  - 初期コードブック $C_0=\{Y_0(i)\}\ (i=1,2,\ldots,N)$
  - トレーニング集合 T を N 個の集合 $R_0(i)$ に分割（最近傍探索）

$$R_0(i)=\{ \boldsymbol{X} \in \boldsymbol{T} : d(\boldsymbol{X},\boldsymbol{Y_0}(i)) < d(\boldsymbol{X},\boldsymbol{Y_0}(j));\ for\ j \neq i \}$$

  - セントロイドベクトルの計算

$$\boldsymbol{Y_1}(i)=E[\boldsymbol{X} \mid \boldsymbol{X} \in R_0(i)]$$

  - 全体での誤差が一定以下になるまで，この処理の繰り返し

Vector Quantization

- Codebook Generation ... LBG Algorithm (Linde, Buzo, Gray)
  - Initial codebook $C_0=\{Y_0(i)\}\ (i=1,2,\ldots,N)$
  - Divide training set T to N sets $R_0(i)$ (Nearest Neighbor Search)

$$R_0(i)=\{ \boldsymbol{X} \in \boldsymbol{T} : d(\boldsymbol{X},\boldsymbol{Y_0}(i)) < d(\boldsymbol{X},\boldsymbol{Y_0}(j));\ for\ j \neq i \}$$

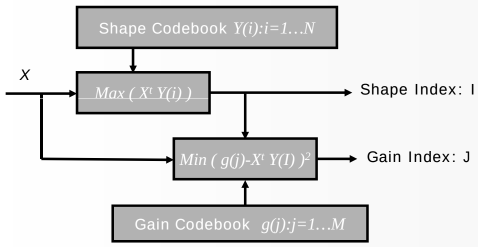
  - Calculation of Centroid Vector

$$\boldsymbol{Y_1}(i)=E[\boldsymbol{X} \mid \boldsymbol{X} \in R_0(i)]$$

  - Repeat this procedure until total error will reach to a certain level

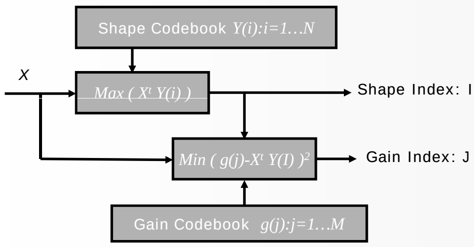
ベクトル量子化 (2)

- Shape-Gain ベクトル量子化



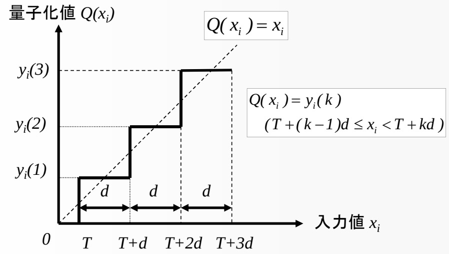
Vector Quantization (2)

- Shape-Gain Vector Quantization



線形量子化

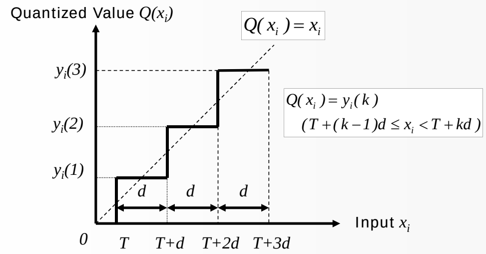
- 量子化間隔 d が一定(一様量子化とも呼ぶ)



信号理論 / Signal Theory 49

Linear Quantization

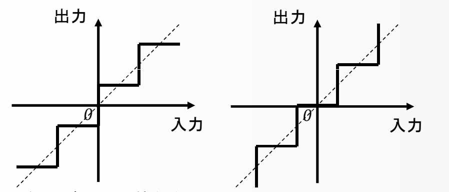
- Quantization Step d is constant (uniform quantization)



信号理論 / Signal Theory 50

線形量子化 (2)

- ミッドライザ型とミッドトレッド型

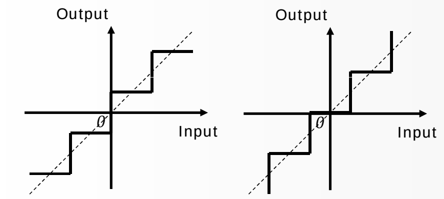


ミッドライザ型 (0近傍を強調)　ミッドトレッド型(0近傍を抑圧)

信号理論 / Signal Theory 51

Linear Quantization (2)

- Mid-riser and Mid-tread Type



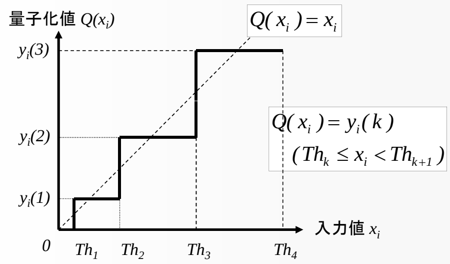
Mid-riser Type
(emphasize around zero)

Mid-tread Type
(suppress around zero)

信号理論 / Signal Theory 52

非線形量子化

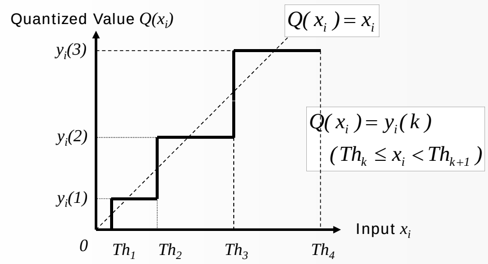
- 量子化ステップが非一様



信号理論 / Signal Theory 53

Non-linear Quantization

- Quantization Step in non-uniform



信号理論 / Signal Theory 54

非線形量子化 (2)

- 信号の特性と量子化ステップの関係
  - 入力信号の出現確率に合わせる
  - 均一な量子化に比べ、平均的な量子化誤差を低減できる
- DPCMのように0付近の信号の出現確率が高い場合
  - 0付近の入力値は出現頻度が高いので細かく量子化
  - 大きな入力値は、出現頻度が低いので粗く量子化
- 入力信号の出現確率と量子化ステップの関係
  - 平均的な量子化誤差の最適化(最小化)が可能

Non-linear Quantization (2)

- Relation of signal characteristics and quantization step
  - Fit to input signal's occurrence probability
  - Can reduce average distortion than linear quantization
- In case of high occurrence probability around zero such as DPCM
  - Fine quantization around zero
  - Coarse quantization to large inputs
- Relation of input's occurrence probability and quantization steps
  - Optimization (Minimization) of the average distortion is possible

Max量子化器

- 信号 x の確率密度 $p(x)$ に関して平均2乗誤差を最小化する量子化器
- しきい値 $t_0, t_1, \dots, t_n$, 量子化値 $q_0, q_1, \dots, q_n$ に対する設計手順
  - Step.1 しきい値の設定

$$t_i = (q_{i-1} + q_i)/2 \quad (i = 1,2,\dots,n-1)$$

ただし $t_0 = -\infty \quad t_n = \infty$

  - Step.2 量子化値の設定(重心)

$$q_i = \int_{t_i}^{t_{i+1}} xp(x)dx \Big/ \int_{t_i}^{t_{i+1}} p(x)dx \quad (i = 0,1,\dots,n-1)$$

Max Quantizer

- Minimize MSE with regard to signal x's probability $p(x)$
- Design procedure to threshold $t_0, t_1, \dots, t_n$ and quantized value $q_0, q_1, \dots, q_n$
  - Step.1 Set thresholds

$$t_i = (q_{i-1} + q_i)/2 \quad (i = 1,2,\dots,n-1)$$

where $t_0 = -\infty \quad t_n = \infty$

  - Step.2 Obtain quantized values (Centroid)

$$q_i = \int_{t_i}^{t_{i+1}} xp(x)dx \Big/ \int_{t_i}^{t_{i+1}} p(x)dx \quad (i = 0,1,\dots,n-1)$$

Max量子化器 (2)

- 量子化誤差(平均2乗誤差)の計算

$$e^2 = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} p(x)(x - q_i)^2 dx$$

- 最適量子化を与えるしきい値は上式を t_i で偏微分して0と置けば

$$t_i = (q_i + q_{i-1})/2 \quad (i = 1,2,\dots,n-1)$$

- 同様に q_i で偏微分すれば量子化値は重心で与えられる(設計手順の理論的背景)
- 確率密度 $p(x)$ が特殊な場合しか解析的に解けず、一般に数値計算で求める必要が生じる

Max Quantizer (2)

- Calculation of Quantization Error (MSE)

$$e^2 = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} p(x)(x - q_i)^2 dx$$

- Optimum threshold is obtained by taking derivatives of t_i

$$t_i = (q_i + q_{i-1})/2 \quad (i = 1,2,\dots,n-1)$$

- Similarly, by taking derivatives of q_i, quantized value is given by centroid (Theoretical background for design procedure)
- Cannot solve analytically for general probability density $p(x)$, normally use numerical calculation is employed

Max量子化器 (3)

- 入力信号の確率密度 $p(x)$ が平均0、分散1の正規分布の場合

$N=4$	しきい値	レベル
i	t_i	q_i
0		-1.510
1	-0.9816	-0.4528
2	0.0	0.4528
3	0.9816	1.510
e^2	0.1175	

$N=8$	しきい値	レベル
i	t_i	q_i
0		-2.152
1	-1.748	-1.344
2	-1.050	-0.7560
3	-0.5006	-0.2451
4	0.0	0.2451
5	0.5006	0.7560
6	1.050	1.344
7	1.748	2.152
e^2	0.03454	

Max Quantizer (3)

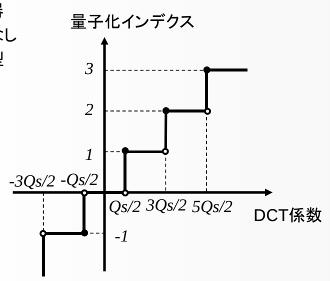
- When input signal's probability $p(x)$ is Normal distribution with average zero and variance 1

$N=4$	threshold	level
i	t_i	q_i
0		-1.510
1	-0.9816	-0.4528
2	0.0	0.4528
3	0.9816	1.510
e^2	0.1175	

$N=8$	threshold	level
i	t_i	q_i
0		-2.152
1	-1.748	-1.344
2	-1.050	-0.7560
3	-0.5006	-0.2451
4	0.0	0.2451
5	0.5006	0.7560
6	1.050	1.344
7	1.748	2.152
e^2	0.03454	

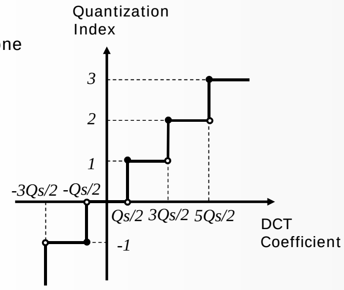
JPEG量子化

- 一様量子化器
- デッドゾーンなし
- ミッドトレッド型



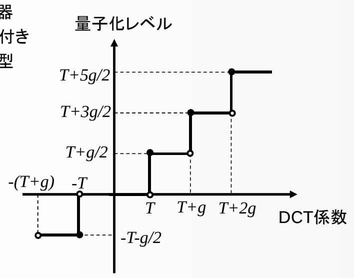
JPEG Quantization

- Uniform
- No dead-zone
- Mid-tread



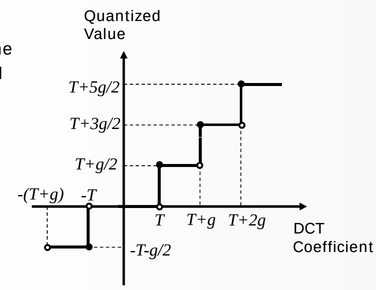
H.261量子化

- 一様量子化器
- デッドゾーン付き
- ミッドトレッド型



H.261 Quantization

- Uniform
- Dead-zone
- Mid-tread



H.261量子化 (2)

- Intra mode（フレーム内符号化モード）
  - DC係数はデッドゾーンなし、量子化ステップ値を8に固定
- Inter mode（フレーム間差分符号化モード）
  - デッドゾーンの例 $T=g$
  - 量子化ステップ $g=2q \quad (q=1, 2, \ldots, 31)$
  - しきい値 $t_i=ig \quad (i= -127, \ldots, -1, 1, \ldots, 127)$
  - 量子化値 $q_i= (t_i+t_{i+1})/2 \quad (i>0)$
    $(t_i+t_{i-1})/2 \quad (i<0)$
  - IDCTミスマッチ処理
    - If(q_i is even) $q_i=q_i-1$;

H.261 Quantization (2)

- Intra mode (Intra-frame coding mode)
  - No dead-zone for DC Coefficient, Quantization Step=8
- Inter mode (Inter-frame difference coding mode)
  - Example of dead-zone $T=g$
  - Quantization step $g=2q \quad (q=1, 2, \ldots, 31)$
  - Threshold $t_i=ig \quad (i= -127, \ldots, -1, 1, \ldots, 127)$
  - Quantization value $q_i= (t_i+t_{i+1})/2 \quad (i>0)$
    $(t_i+t_{i-1})/2 \quad (i<0)$
  - IDCT mismatch process
    - If(q_i is even) $q_i=q_i-1$;

H.261量子化 (3)

- IDCTミスマッチ処理

q	g	t_i (=ig)					q_i			
		$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=1$	$i=2$	$i=3$	$i=4$
1	2	2	4	6	8	10	3	5	7	9
2	4	4	8	12	16	20	6→5	10→9	14→13	18→17
3	6	6	12	18	24	30	9	15	21	27
4	8	8	16	24	32	40	12→11	20→19	28→27	36→35

H.261 Quantization (3)

- IDCT Mismatch Process

q	g	t_i (=ig)					q_i			
		$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=1$	$i=2$	$i=3$	$i=4$
1	2	2	4	6	8	10	3	5	7	9
2	4	4	8	12	16	20	6→5	10→9	14→13	18→17
3	6	6	12	18	24	30	9	15	21	27
4	8	8	16	24	32	40	12→11	20→19	28→27	36→35

MPEG-1量子化

- Intra mode（フレーム内符号化モード）
  - DC係数はデッドゾーンなし、量子化ステップ値を8に固定
  - AC係数はJPEGと同様
  - デッドゾーンなし、一様量子化器
  - 量子化マトリクス使用
- Inter mode（フレーム間差分符号化モード）
  - DC係数、AC係数ともにH.261と同様
  - デッドゾーン付き一様量子化器
  - 量子化マトリクスを使用

MPEG-1 Quantization

- Intra mode (Intra-frame coding mode)
  - DC coeff: no dead-zone, quantization step=8
  - AC coeff is the same as JPEG
  - No dead-zone, uniform quantizer
  - Quantization matrix
- Inter mode (Inter-frame difference coding mode)
  - DC coeffs and AC coeffs are the same as H.261
  - Dead-zone, uniform quantizer
  - Quantization matrix

MPEG-1量子化 (2)

- 量子化マトリクスを用いた量子化処理
  - Intra, AC係数の量子化

$$A(u,v)=\frac{16X(u,v)\pm W_{Intra}(u,v)/2}{W_{Intra}(u,v)}$$
$$N(u,v)=\frac{A(u,v)\pm q}{2q}$$
$$Q_{AC-Intra}(u,v)=2qN(u,v)W_{intra}(u,v)/16$$

MPEG-1 Quantization (2)

- Quantization using Matrix
  - Quantization to Intra, AC Coefficients

$$A(u,v)=\frac{16X(u,v)\pm W_{Intra}(u,v)/2}{W_{Intra}(u,v)}$$
$$N(u,v)=\frac{A(u,v)\pm q}{2q}$$
$$Q_{AC-Intra}(u,v)=2qN(u,v)W_{intra}(u,v)/16$$

MPEG-1量子化 (3)

- 量子化マトリクスを用いた量子化処理
  - Inter, DC+AC係数の量子化

$$A(u,v)=\frac{16X(u,v)\pm W_{Inter}(u,v)/2}{W_{Inter}(u,v)}$$
$$N(u,v)=\frac{A(u,v)}{2q}\quad (q:odd)$$
$$N(u,v)=\frac{A(u,v)\pm 1}{2q}\quad (q:even)$$
$$Q_{Inter}(u,v)=(2N(u,v)\pm 1)qW_{inter}(u,v)/16$$
$$Q_{Inter}(u,v)=Q_{Inter}(u,v)\mp 1\quad (if\ Q_{Inter}(u,v):even)$$

MPEG-1 Quantization (3)

- Quantization using Matrix
  - Quantization to Inter, DC+AC Coefficients

$$A(u,v)=\frac{16X(u,v)\pm W_{Inter}(u,v)/2}{W_{Inter}(u,v)}$$
$$N(u,v)=\frac{A(u,v)}{2q}\quad (q:odd)$$
$$N(u,v)=\frac{A(u,v)\pm 1}{2q}\quad (q:even)$$
$$Q_{Inter}(u,v)=(2N(u,v)\pm 1)qW_{inter}(u,v)/16$$
$$Q_{Inter}(u,v)=Q_{Inter}(u,v)\mp 1\quad (if\ Q_{Inter}(u,v):even)$$

MPEG-1量子化 (4)

- 量子化マトリクス

$$\begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix} \quad \begin{bmatrix} 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \end{bmatrix}$$

$W_{intra}(u,v)$ $\quad$ $W_{inter}(u,v)$

MPEG-1 Quantization (4)

- Quantization Matrix

$$\begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix} \quad \begin{bmatrix} 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \end{bmatrix}$$

$W_{intra}(u,v)$ $\quad$ $W_{inter}(u,v)$

MPEG-2量子化

- Intra mode（フレーム内符号化モード）
  - DC係数はデッドゾーンなし、量子化ステップ値を8に固定
  - デッドゾーンなし、一様量子化器
  - 量子化マトリクス使用
- Inter mode（フレーム間差分符号化モード）
  - デッドゾーン付き一様量子化器
  - 量子化マトリクスを使用

MPEG-2 Quantization

- Intra mode (Intra-frame coding mode)
  - DC coeff: no dead-zone, quantization step=8
  - No dead-zone, uniform quantizer
  - Quantization matrix
- Inter mode (Inter-frame difference coding mode)
  - With dead-zone, uniform quantizer
  - Quantization matrix

MPEG-2量子化 (2)

- 量子化マトリクスを用いた量子化処理
  - Intra, AC係数の量子化

$$A(u,v)=\frac{16X(u,v)+W_{Intra}(u,v)/2}{W_{Intra}(u,v)}$$
$$N(u,v)=\frac{A(u,v)\pm aq/b}{2q}$$
$$Q_{Intra-AC}(u,v)=2qN(u,v)W_{Intra}(u,v)/16$$

MPEG-2 Quantization (2)

- Quantization by Matrix
  - Quantization to Intra, AC Coefficients

$$A(u,v)=\frac{16X(u,v)+W_{Intra}(u,v)/2}{W_{Intra}(u,v)}$$
$$N(u,v)=\frac{A(u,v)\pm aq/b}{2q}$$
$$Q_{Intra-AC}(u,v)=2qN(u,v)W_{Intra}(u,v)/16$$

MPEG-2量子化 (3)

- 量子化マトリクスを用いた量子化処理
  - Inter, DC+AC係数の量子化

$$A(u,v)=\frac{16X(u,v)+W_{Inter}(u,v)/2}{W_{Inter}(u,v)}$$
$$N(u,v)=\frac{A(u,v)}{2q}\quad(q:odd)$$
$$N(u,v)=\frac{A(u,v)\pm 1}{2q}\quad(q:even)$$
$$Q_{Inter}(u,v)=(2N(u,v)\pm 1)qW_{Inter}(u,v)/16$$

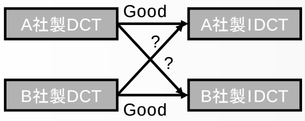
MPEG-2 Quantization (3)

- Quantization by Matrix
  - Quantization to Inter, DC+AC Coefficients

$$A(u,v)=\frac{16X(u,v)+W_{Inter}(u,v)/2}{W_{Inter}(u,v)}$$
$$N(u,v)=\frac{A(u,v)}{2q}\quad(q:odd)$$
$$N(u,v)=\frac{A(u,v)\pm 1}{2q}\quad(q:even)$$
$$Q_{Inter}(u,v)=(2N(u,v)\pm 1)qW_{Inter}(u,v)/16$$

IDCTミスマッチ

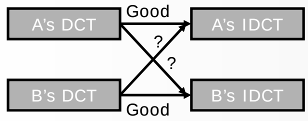
- 異なる演算精度でIDCTが実行された場合に起こる



- 8x8 DCTの計算式
  - 先頭に1/4を掛ける部分が存在する
  - 4k+2(k:整数)なるレベルが存在すると問題

IDCT Mismatch

- Happens when IDCT is operated different accuracy



- 8x8 DCT operation
  - Multiply 1/4 at the head of operation
  - 4k+2(k:integer) level causes problem

IDCTミスマッチ (2)

- 8x8 IDCTにおける積和が4k+2の場合

$$\begin{aligned} x(n,m) &= \frac{1}{4}\left(\sum_{u=0}^{7}\sum_{v=0}^{7} C(u)C(v)X(u,v)\cos\frac{(2n+1)u\pi}{16}\cos\frac{(2m+1)v\pi}{16}\right) \\ &= \frac{1}{4}(4k+2) \\ &= k+0.5 \end{aligned}$$

- 演算手法，演算精度により結果が異なる

$$\begin{aligned} x(n,m) &= k+0.500000000\ldots \quad (k+1) \\ &or \\ &= k+0.499999999\ldots \quad (k) \end{aligned}$$

IDCT Mismatch(2)

- For Multiplication and addition to 4k+2 at 8x8 IDCT

$$\begin{aligned} x(n,m) &= \frac{1}{4}\left(\sum_{u=0}^{7}\sum_{v=0}^{7} C(u)C(v)X(u,v)\cos\frac{(2n+1)u\pi}{16}\cos\frac{(2m+1)v\pi}{16}\right) \\ &= \frac{1}{4}(4k+2) \\ &= k+0.5 \end{aligned}$$

- Result varies by operation algorithm and accuracy

$$\begin{aligned} x(n,m) &= k+0.500000000\ldots \quad (k+1) \\ &or \\ &= k+0.499999999\ldots \quad (k) \end{aligned}$$

IDCTミスマッチ対策

- 量子化値の制限
  - H.261… 偶数値の奇数化
  - MPEG-1… 偶数値の奇数化(0方向に1近づける)
    逆量子化の前に処理
  - MPEG-2… 8x8DCT係数の最高周波数の係数のLSBを反転
    逆量子化の後に処理
- IDCT精度の規定
  - IEEE 1180 (Dec. 1990)

IDCT Mismatch Remedy

- Restriction to Quantized Value
  - H.261… Oddification of even value
  - MPEG-1… Oddification of even value (decrease 1 towards 0)
    Process before inverse quantization
  - MPEG-2… reverse LSB of a coefficient at highest frequency in 8x8DCT coefficients
    Process before inverse quantization
- Standard for IDCT Accuracy
  - IEEE 1180 (Dec. 1990)