Quantization

AD converters

Two steps.

1) Sampling: result, sequences of real numbers.
2) Quantization: result, the samples represented by finite many bits.

Uniform quantization

Thumb rule

1) The signal is bounded: $\exists\, K > 0$, such that $-K \leq x[k] \leq K$.
2) The smallest value that is considered as non-zero: $m > 0$.
3) Uniform partition of the range, dynamics of the signal, required number of levels: between $\dfrac{2K}{m} + 1$ and $\dfrac{2K}{m} + 3$.
4) Number of bits needed: $N \in \mathbb{N}$, where $2^{N-2} < \dfrac{K}{m} + 1 < 2^{N-1}$.

Coding

First bit is the sign of $x[k]$.

Use $N-1$ to code $|x[k]|$.

In order to do this find $j = 0, \ldots, 2^{n-1}-1$ for which $\left| |x[k]| - \frac{j}{m} \right|$ is the smallest.

Take the binary form of j.

Example

The range of voltage of the signal: -180.0mV – 180.0mV.

The smallest sensible value: 3mV.

Dynamics of the signal: 120.

Bits: $2^5 < 180/3 = 60 < 2^6$. 7 bits are needed, including the sign bit.

Measured value: 17.6mV. $5 \cdot 3 < 17.6 < 6 \cdot 3$. $j = 6$.

Code: 0000110. The first 0 means positive sign.

Companders

Telephone signal

Dynamic range in telephone technology: approx. $16386 = 2^{14}$.
$10 \log_{10} 16386 \approx 42$ decibel.

It is the ratio between the quietest and loudest sound that can be represented in the signal.

Uniform quantization: the number of bits is 14.

Why compression is possible?

The signal is to be perceived as audio by a human.

The point: the perceived acoustic intensity level or loudness is logarithmic.

The ear is more sensitive in the quiet than in the loud range of sound.

Telephone signal compression standard in North America and Japan: μ-law.

Compressor

Reduction of 14 bits to 8 bit.

The steps.

1) The signal is uniformly quantized by using 14 bits into the range $-8192, \ldots, 0, \ldots, 8191 : | \; E \; | \; X_{12} \; | \quad \ldots \quad | \; X_0 \; |$, where E is the sign bit.
2) The value $X = \pm \sum_{k=0}^{12} X_k 2^k$ is linearly transformed into the interval $[-1, 1] : x = \dfrac{X}{2^{13}}$.
3) Apply the μ-transform: $y = \operatorname{sgn} x \, \dfrac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}$, where $\mu = 255$. Note that $y \in [-1, 1]$.
4) Multiplication of y by $2^7 = 128$. The result is in the range $-128, \ldots, 0, \ldots, 127$. Round the value to the nearest integer Y, and express it by 8 bits.

Examples

Example 1:

1. step: $X = 2147$. $X = 0100001100011$

2. step: $x = \dfrac{2147}{8192} = 0.262084...$

3. step: $y = \operatorname{sgn} x \, \dfrac{\ln(1+\mu|x|)}{\ln(1+\mu)} = 0.760485...$

4. step: $128 \cdot y = 97.342192...,\ Y = 98,\ Y = 01100010$

Examples

Example 1:

1. step: $X = 2147$. $X = 0100001100011$

2. step: $x = \dfrac{2147}{8192} = 0.262084...$

3. step: $y = \operatorname{sgn} x \, \dfrac{\ln(1+\mu|x|)}{\ln(1+\mu)} = 0.760485...$

4. step: $128 \cdot y = 97.342192...$, $Y = 98$, $Y = 01100010$

Example 2.

1. step: $X = 6355$. $X = 01100011010011?$

2. step: $x = \dfrac{6355}{8192} = 0.775756...$

3. step: $y = \operatorname{sgn} x \, \dfrac{\ln(1+\mu|x|)}{\ln(1+\mu)} = 0.954413...$

4. step: $128 \cdot y = 122.164872$, $Y = 122$, $Y = 01111010$

Expander

The steps.

1) Divide Y by 128 to get $\widetilde{y} \in [-1, 1]$.

2) Apply the inverse μ-transform:

$$\widetilde{x} = \operatorname{sgn} \widetilde{y} \, \frac{(1+\mu)^{|\widetilde{y}|} - 1}{\mu} = \operatorname{sgn} \widetilde{y} \, \frac{256^{|\widetilde{y}|} - 1}{255} \in [-1, 1].$$

3) Multiplication of $\widetilde{x}$ by $2^{13} = 8192$. The result is in the range $-8192, \ldots, 0, \ldots, 8191$. Round the value to the nearest integer $\widetilde{X}$, and express it by 14 bits.

Expander

The steps.

1) Divide Y by 128 to get $\widetilde{y} \in [-1, 1]$.
2) Apply the inverse μ-transform:
$$\widetilde{x} = \operatorname{sgn} \widetilde{y} \, \frac{(1+\mu)^{|\widetilde{y}|} - 1}{\mu} = \operatorname{sgn} \widetilde{y} \, \frac{256^{|\widetilde{y}|} - 1}{255} \in [-1, 1].$$
3) Multiplication of $\widetilde{x}$ by $2^{13} = 8192$. The result is in the range $-8192, \ldots, 0, \ldots, 8191$. Round the value to the nearest integer $\widetilde{X}$, and express it by 14 bits.

Examples

Example 1: $Y = 98$, $\widetilde{y} = 0.765625$, $\widetilde{x} = 0.268720...$,
$\widetilde{x} \cdot 8192 = 2201.3...$, $\widetilde{X} = 2201$. Recall: $X = 2147$.

Relative error
$$\frac{|2147 - 2201|}{2147} = 0.0251... \approx 2.5\%$$

Examples (cont.)

Example 2: $Y = 122$, $\widetilde{y} = 0.953125$, $\widetilde{x} = 0.770207...$, $\widetilde{x} \cdot 8192 = 6309.5...$, $\widetilde{X} = 6310$. Recall: $X = 6355$

Relative error

$$\frac{|6310 - 6355|}{6355} = 0,0070... \approx 0.7\%$$

Practical realization

The steps

1) Addition 33 to $|X|$: $Z = |X| + 33$.
Keep the sign bit E.

2) Take the binary form of Z with 13 bits.
$| Z_{12} | Z_{11} | \dots |Z_1| , |Z_0|$

3) The position of the leading 1 digit is: between 5 and 12.
Position: $S + 5$, where $S = 0, \dots, 7$.
Code S by using 3 digits: $| S_2 | S_1 | S_0 |$. $S = 4S_2 + 2S_1 + S_0$.

4) Keep the 4 digits $| Z_{S+4} | Z_{S+3} | Z_{S+2} | Z_{S+1} |$ in Z that follow the leading 1 digit.

5) The code:

$$| E | S_2 | S_1 | S_0 | Z_{S+4} | Z_{S+3} | Z_{S+2} | Z_{S+1} |$$
$$=$$
$$| A_7 | A_6 | A_5 | A_4 | A_3 | A_2 | A_1 | A_0 |$$

Decoding

$$| A_7 | A_6 | A_5 | A_4 | A_3 | A_2 | A_1 | A_0 |$$
$$=$$
$$| E | S_2 | S_1 | S_0 | Z_{S+4} | Z_{S+3} | Z_{S+2} | Z_{S+1} |$$

Reconstruction of $S = 4A_6 + 2A_5 + A_4 + 5$.

Then

$$\begin{aligned} |\widetilde{X}| &= 2^{S+5} + A_3 2^{S+4} + A_2 2^{S+3} + A_1 2^{S+2} + A_0 2^{S+1} + 2^S - 33 \\ &= 2^S(2^5 + A_3 2^4 + A_2 2^3 + A_1 2^2 + A_0 2^1 + 1) - 33\,. \end{aligned}$$

Examples

Example 1:

Coding

$X = 2147$, $Z = 2180$, $Z = 0100010000100$.

Position of leading 1 digit: 11.

Then $S = 6 = 1 \cdot 4 + 1 \cdot 2 + 0 \cdot 1$.

The S digit will be followed by 0001. The code without the sign bit:

$$|\,1\,|\,1\,|\,0\,|\,0\,|\,0\,|\,0\,|\,1\,|.$$

Decoding:

$$\widetilde{Z} = 2^6(2^5 + 0 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1) = 2240.$$

$2240 = 0100011000000$.

Then $\widetilde{X} = 2207$. Relative error: $\dfrac{2207 - 2147}{2147} = 0,0279... < 2.8\%$.

Examples

Example 1:

Coding

$X = 6355$, $Z = 6388$, $Z = 1100011110100$.

Position of leading 1 digit: 12.

Then $S = 7 = 1 \cdot 4 + 1 \cdot 2 + 1 \cdot 1$.

The S digit will be followed by 1000. The code without the sign bit:

$$|\,1\,|\,1\,|\,1\,|\,1\,|\,0\,|\,0\,|\,0\,|.$$

Decoding:

$$\widetilde{Z} = 2^7(2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1) = 6272.$$

$6272 = 1100010000000$.

Then $\widetilde{X} = 6239$. Relative error: $\dfrac{6355 - 6239}{6355} = 0,0182... \approx 1.8\%$.

General relative error

The relative error with respect to Z: $\dfrac{|\widetilde{Z} - Z|}{Z} < \dfrac{2^s}{2^{s+5}} = \dfrac{1}{32} = 3,125\%$

Remarks

The role of 33 :
a) Takes the leading 1 to positions between 5 and 7. 32 would do this job.
b) Check the case $X = 0$. Using 32 we get $\widetilde{X} = 1$.

The role of the term 2^S in the decoding: Decreasing the error.

A-law (Europe):

$$y = \begin{cases} \operatorname{sgn} x \, \dfrac{A|x|}{1 + \ln A} & 0 \leq |x| < \frac{1}{A}, \\[2ex] \operatorname{sgn} x \, \dfrac{1 + \ln A|x|}{1 + \ln A} & \frac{1}{A} \leq |x| < 1, \end{cases}$$

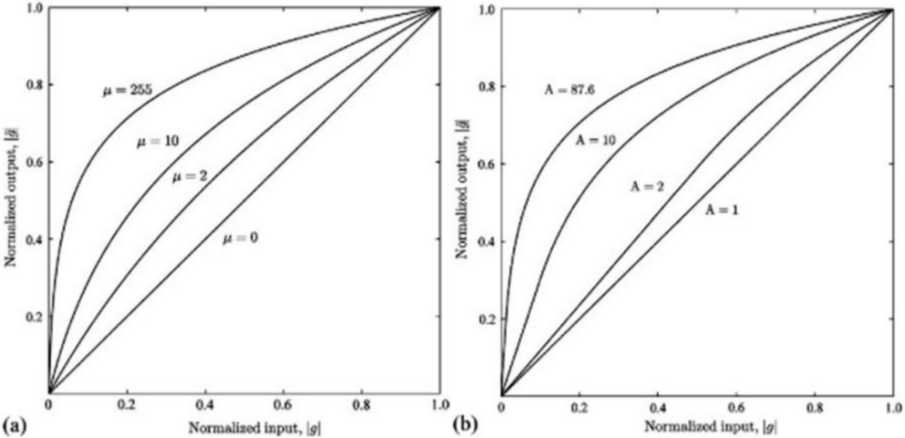

Figure: The graphs of μ-law and A-law functions