

高斯信号的量化分析*

刘小汇¹, 黄龙¹, 谢金石²

(1. 国防科技大学 电子科学与工程学院, 湖南 长沙 410073; 2. 北京环球信息应用开发中心, 北京 100094)

摘要: 由于器件的约束条件不同, 在数字信号处理中通常要对 AD 采样后的信号进行量化处理, 如何提高量化的性能是量化的关键。针对 $N(\mu, \sigma^2)$ 高斯分布的量化信号, 使用理论推导的方式, 得到量化损耗与限幅门限的关系式, 并推导出对应量化损耗最小的最优限幅门限公式, 其推导方法同样适用于其他随机信号。仿真试验证明了该公式的正确性, 结论可应用于工程实现。

关键词: 高斯信号; 量化; 限幅门限; 量化损耗

中图分类号: TN967.1 **文献标识码:** A

Analysis of Quantization of Gaussian Signals

LIU Xiao-hui¹, HUANG Long¹, XIE Jin-shi²

(1. College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073, China;

2. Beijing Global Information Application and Development Center, Beijing 100094, China)

Abstract: Quantization will be necessary in the digital signal processing because of the hardware's resource. It is important to improve the performance of the quantizer. An expression of quantization distortion to clip level was derived theoretical for $N(\mu, \sigma^2)$ Gaussian pdf signal. Also the expression of the minimal quantization distortion and optimum clip level was deducted. The result of computer simulation is that the expression can be used to estimate the optimum clip level effectively and exactly in the related applications.

Key words: Gaussian signal; quantization; clip level; quantization distortions

数字信号处理中, AD 是信号处理的前提, 而量化又是 AD 转换的关键环节, 如何获取尽可能高的性能是对量化研究的关键。对于均匀量化, 目前的研究多集中在不同的数字信号处理过程中, 量化位数对系统性能的影响上^[1-3]; 在量化门限对量化性能的影响方面, 许多研究指出, 通过合理选择量化的门限可以得到最优的量化损耗^[4-6], 然而以上文献只通过数值仿真的方法给出了结论性的说明; 文献[7]分析了量化间隔与量化信噪比损耗的关系, 但是没有考虑限幅效应对损耗的影响; Max 在 1960 年计算出了 $N(0, 1)$ 标准高斯分布的量化信号的最优量化步长与最小均方误差的对应关系表^[8], 但所给出的公式无法得到解析的解。本文以更普遍的 $N(\mu, \sigma^2)$ 高斯分布的信号为对象, 首先分析了量化损耗的组成, 得出限幅门限对量化性能的影响公式, 然后推导出最小量化损耗与限幅门限的关系式, 说明量化信号的信噪比系数 (μ/σ) 对量化损耗有影响, 并

得到不同信噪比系数下最小量化损耗与对应限幅门限的关系表, 最后通过实际信号仿真表明公式的正确性。

1 量化信号的损耗分析

对 AD 采样后的离散信号 $x(n)$ 进行 $B+1$ (最高位为符号位) 比特均匀量化, 量化间隔为 Δ , 量化结果 $\hat{x}(n)$:

$$\hat{x}(n) = \begin{cases} (2^B - 1)\Delta & x(n) \geq (2^B - 3/2)\Delta \\ i\Delta & (i - 1/2)\Delta \leq x(n) < (i + 1/2)\Delta \\ -2^B\Delta & x(n) < -(2^B - 1/2)\Delta \end{cases} \quad (1)$$

其中, $i = -(2^B - 1), \dots, -1, 0, 1, \dots, (2^B - 2)$, 总的量化级数为 $L = 2^{B+1}$, 设量化满量程值 X_m , 则量化间隔 $\Delta = X_m/2^{B+1}$ 。

量化会引入量化噪声, 量化噪声会带来一定的信噪比损耗, 称量化损耗。若量化信号的噪声

* 收稿日期: 2011-03-02

作者简介: 刘小汇(1976—), 女, 副研究员, 博士生。

方差为 σ^2 , 由量化引起的量化噪声方差为 σ_q^2 , 假设量化信号与量化噪声不相关, 则量化后总的噪声方差: $\sigma_\Sigma^2 = \sigma^2 + \sigma_q^2$, 定义量化损耗^[2]

$$Loss = \frac{\sigma_\Sigma^2}{\sigma^2} = 1 + \frac{\sigma_q^2}{\sigma^2} \quad (2)$$

量化信号的概率分布不同, 其量化损耗会有不同。不失一般性, 本文分析高斯分布的量化信号, 其他信号的分析方法类似。假设 $x(n)$ 为 $N(\mu, \sigma^2)$ 分布的高斯信号, 概率密度函数为 $f(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} / (\sigma \sqrt{2\pi})$ 。当量化级数为 L 时, 量化区间为 $Q_1 = (-\infty, X_1), Q_2 = (X_2, X_3), \dots, Q_L = (X_{L-1}, +\infty)$, 则量化误差 $\varepsilon(i) = x(i) - \hat{x}(i)$, 其中 $i = -2^B, \dots, 0, \dots, (2^B - 1)$ 。

易得 $E[\varepsilon] = 0$, 量化噪声方差 σ_q^2 定义为

$$\begin{aligned} \sigma_q^2 &= \sum_{i=1}^L \int_{Q_i} \varepsilon^2 f(\varepsilon) d\varepsilon \\ &= \int_{-\infty}^{X_1} \varepsilon^2 f(\varepsilon) d\varepsilon + \sum_{i=2}^{L-1} \int_{Q_i} \varepsilon^2 f(\varepsilon) d\varepsilon + \int_{X_{L-1}}^{+\infty} \varepsilon^2 f(\varepsilon) d\varepsilon \end{aligned} \quad (3)$$

其中, X_1 和 X_{L-1} 分别为负、正界的量化满量程值, 对于 $B+1$ 比特(最高位为符号位)量化的噪声方差, 由式(3)可知 $\sigma_q^2 = \sigma_c^2 + \sigma_e^2$, 其中 $\sigma_e^2 = \sum_{i=2}^{L-1} \int_{Q_i} \varepsilon^2 f(\varepsilon) d\varepsilon$ 称为量化间隔误差, $\sigma_c^2 = \int_{-\infty}^{X_1} \varepsilon^2 f(\varepsilon) d\varepsilon + \int_{X_{L-1}}^{+\infty} \varepsilon^2 f(\varepsilon) d\varepsilon$ 称为限幅误差。

文献[9]给出 $N(0, \sigma^2)$ 的高斯分布信号量化间隔误差的公式, 将公式进行扩展, 可以推导得到服从 $N(\mu, \sigma^2)$ 高斯分布的量化信号, 其量化间隔误差的公式:

$$\begin{aligned} \sigma_e^2 &= \frac{\Delta^2}{12} \left[1 + \frac{12}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos\left(\frac{2\pi n \mu}{\Delta}\right) \right. \\ &\quad \left. \exp\left(-\frac{2\pi^2 n^2 \sigma^2}{\Delta^2}\right) \right] \end{aligned} \quad (4)$$

当 $\sigma/\Delta \geq 1$ 时, 误差可以近似为

$$\sigma_e^2 \approx \frac{\Delta^2}{12} \quad (5)$$

为了简化计算, 对于多比特量化, 假设限幅值 $|X_1| = |X_{L-1}| = X_m$, 量化门限 $(2^B - 1)\Delta \approx 2^B \Delta = A$, 限幅的误差定义为

$$\varepsilon(i) = x(i) - \hat{x}(i) = \begin{cases} x(i) - A & x(i) > X_m \\ 0 & |x(i)| \leq X_m \\ x(i) + A & x(i) < -X_m \end{cases} \quad (6)$$

其中

$$X_m = \frac{A}{1 - 2^{-B}} = 2^B \Delta \quad (7)$$

将 x 正限幅时的概率分布函数记做 P_{c+} , 负限幅时的概率分布函数记做 P_{c-} , 它们分别可以表示为

$$\begin{aligned} P_{c+} &= P(x > X_m) \\ &= \frac{1}{\sigma \sqrt{2\pi}} \int_{X_m}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{X_m - \mu}{\sigma \sqrt{2}}\right) \end{aligned} \quad (8)$$

$$\begin{aligned} P_{c-} &= P(x < -X_m) \\ &= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{-X_m} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{X_m + \mu}{\sigma \sqrt{2}}\right) \end{aligned} \quad (9)$$

其中, $Q(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ 。

可以计算

$$E[\varepsilon] = \int_{-\infty}^{-X_m} \varepsilon f(\varepsilon) d\varepsilon + \int_{X_m}^{+\infty} \varepsilon f(\varepsilon) d\varepsilon = 0 \quad (10)$$

通过求正限幅时的条件概率函数 $f(x|x > A)$ 和负限幅时的条件概率函数 $f(x|x < -A)$, 可以得到限幅误差 ε 的条件二阶矩, 进而得到限幅误差的表达式:

$$\begin{aligned} \sigma_c^2 &= E[\varepsilon^2] = \int_{-\infty}^{-X_m} \varepsilon^2 f(\varepsilon) d\varepsilon + \int_{X_m}^{+\infty} \varepsilon^2 f(\varepsilon) d\varepsilon \\ &= P_{c+} E[\varepsilon^2 | x > X_m] + P_{c-} E[\varepsilon^2 | x < -X_m] \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{X_m - \mu}{\sigma \sqrt{2}}\right) [\sigma^2 + (A - \mu)^2] \\ &\quad + \frac{\sigma(\mu + X_m - 2A)}{\sqrt{2\pi}} e^{-\frac{(X_m - \mu)^2}{2\sigma^2}} \\ &\quad + \frac{1}{2} \operatorname{erfc}\left(\frac{X_m + \mu}{\sigma \sqrt{2}}\right) [\sigma^2 + (A + \mu)^2] \\ &\quad + \frac{\sigma(X_m - \mu - 2A)}{\sqrt{2\pi}} e^{-\frac{(X_m + \mu)^2}{2\sigma^2}} \end{aligned} \quad (11)$$

由式(5)和式(11)可得到量化噪声方差的表达式:

$$\begin{aligned} \sigma_q^2 &= \sigma_e^2 + \sigma_c^2 \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{X_m - \mu}{\sigma \sqrt{2}}\right) [\sigma^2 + (A - \mu)^2] \\ &\quad + \frac{\sigma(\mu + X_m - 2A)}{\sqrt{2\pi}} e^{-\frac{(X_m - \mu)^2}{2\sigma^2}} \\ &\quad + \frac{1}{2} \operatorname{erfc}\left(\frac{X_m + \mu}{\sigma \sqrt{2}}\right) [\sigma^2 + (A + \mu)^2] \\ &\quad + \frac{\sigma(X_m - \mu - 2A)}{\sqrt{2\pi}} e^{-\frac{(X_m + \mu)^2}{2\sigma^2}} \end{aligned}$$

$$+ \frac{A^2}{12(2^B - 1)^2} \quad (12)$$

对于多比特量化,假设 $X_m \approx A$, 则

$$\begin{aligned} \sigma_q^2 = & \frac{1}{2} \operatorname{erfc}\left(\frac{A-\mu}{\sigma\sqrt{2}}\right) [\sigma^2 + (A-\mu)^2] \\ & + \frac{\sigma(\mu-A)}{\sqrt{2\pi}} e^{-\frac{(A-\mu)^2}{2\sigma^2}} \\ & + \frac{1}{2} \operatorname{erfc}\left(\frac{A+\mu}{\sigma\sqrt{2}}\right) [\sigma^2 + (A+\mu)^2] \\ & - \frac{\sigma(\mu+A)}{\sqrt{2\pi}} e^{-\frac{(A+\mu)^2}{2\sigma^2}} + \frac{A^2}{12(2^B - 1)^2} \quad (13) \end{aligned}$$

于是量化损耗为

$$\begin{aligned} Loss = & \frac{1}{2} \operatorname{erfc}\left(\frac{C-K}{\sqrt{2}}\right) [1 + (C-K)^2] \\ & + \frac{(K-C)}{\sqrt{2\pi}} e^{-\frac{(C-K)^2}{2}} \\ & + \frac{1}{2} \operatorname{erfc}\left(\frac{C+K}{\sqrt{2}}\right) [1 + (C+K)^2] \\ & - \frac{K+C}{\sqrt{2\pi}} e^{-\frac{(C+K)^2}{2}} + \frac{C^2}{12(2^B - 1)^2} + 1 \quad (14) \end{aligned}$$

其中, $A/\sigma = C$ 为限幅门限系数, $\mu/\sigma = K$ 为信噪比系数。由式(14)得到的限幅门限系数与量化损耗之间的关系如图 1 所示 ($K=0$), 相同的信号经过不同量化位数的量化后, 其损耗与限幅门限的关系表明: 量化位数越高, 损耗越小; 存在一个最优的限幅门限, 使得在量化位数一定时, 量化损耗最小; 对于多比特量化, 当限幅门限小于最优门限时, 门限的减小会带来损耗的显著恶化, 当限幅门限大于最优门限, 门限的变化对损耗的影响很

$$- \frac{2\sigma}{\sqrt{2\pi}} [e^{-\frac{(A-\mu)^2}{2\sigma^2}} + e^{-\frac{(A+\mu)^2}{2\sigma^2}}] + (A-\mu) \operatorname{erfc}\left(\frac{A-\mu}{\sigma\sqrt{2}}\right) + (A+\mu) \operatorname{erfc}\left(\frac{A+\mu}{\sigma\sqrt{2}}\right) + \frac{A}{6(2^B - 1)^2} = 0 \quad (17)$$

令 $A/\sigma = C, \mu/\sigma = K$, 整理后得

$$B = \log_2 \left\{ \sqrt{\frac{C}{6 \left[\sqrt{\frac{2}{\pi}} (e^{-\frac{(C-K)^2}{2}} + e^{-\frac{(C+K)^2}{2}}) - (C-K) \operatorname{erfc}\left(\frac{C-K}{\sqrt{2}}\right) - (C+K) \operatorname{erfc}\left(\frac{C+K}{\sqrt{2}}\right) \right] + 1}} \right\} \quad (18)$$

上式即为最小量化损耗下, $N(\mu, \sigma^2)$ 的高斯分布信号 x 的量化位数 B 与限幅门限系数 C 之间的关系式。若定义最小量化损耗对应的量化限幅门限系数为最优系数 C_{opt} , 由公式可见, 最佳限幅门限系数 C_{opt} 的取值与信噪比系数 K 和量化位数 B 均有关系。图 2 是在相同的量化位数(4 比特量化)下, 不同信噪比系数的信号, 其限幅门限系数 C 与量化损耗 $Loss$ 的关系。由图 2 可见, 3 组信号由于信噪比系数 K 值的影响, 其最小量化损耗会有不同, K 值越大, 其对应的损耗也越

小。值得一提的是, 在以上的公式推导中, 由于使用了限幅值 $|X_1| = |X_{L-1}| = X_m$ 以及量化门限 $(2^B - 1)\Delta \approx 2^B \Delta = A$ 等近似处理, 对于低比特(1.5 或 2 比特)量化的情形会带来较大误差, 因此本文的公式对于低比特量化不再适用。

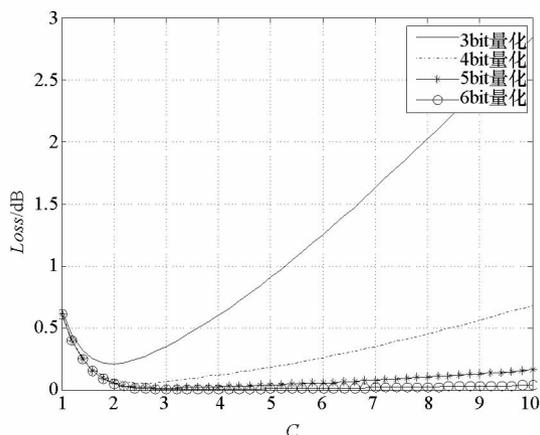


图 1 多比特量化时限幅门限系数与量化损耗的关系
Fig. 1 Quantization distortion vs. clip level under multi-bit quantization

2 最佳限幅门限

以下讨论最佳限幅门限的取值, 通过对式(13)求导可得最佳限幅门限的表达式

$$\frac{\partial \sigma_q^2}{\partial A} = 0 \quad (15)$$

已知误差函数求导数为

$$\frac{d}{dc} \operatorname{erfc}\left(\frac{c}{\sqrt{2}}\right) = -\sqrt{\frac{2}{\pi}} e^{-\frac{c^2}{2}} \quad (16)$$

式(15)可以计算得到

$$- \frac{2\sigma}{\sqrt{2\pi}} [e^{-\frac{(A-\mu)^2}{2\sigma^2}} + e^{-\frac{(A+\mu)^2}{2\sigma^2}}] + (A-\mu) \operatorname{erfc}\left(\frac{A-\mu}{\sigma\sqrt{2}}\right) + (A+\mu) \operatorname{erfc}\left(\frac{A+\mu}{\sigma\sqrt{2}}\right) + \frac{A}{6(2^B - 1)^2} = 0 \quad (17)$$

大, 相应限幅门限也越大。图 3 给出了在 3、4、5、6 比特量化时, 不同 K 值对应的最小量化损耗, 随着量化位数的增大, K 值对量化损耗的影响将越来越小, 表 1 列出了 $B=5$ (6 比特量化) 的量化位数下, 对不同 K 值信号进行量化, 最佳限幅门限与最小量化损耗的结果, 当 K 由 0 增大至 1 时, 带来的损耗只增加了 0.002dB 左右, 可以忽略。作为对比, 同时也列出了数值仿真的结果。

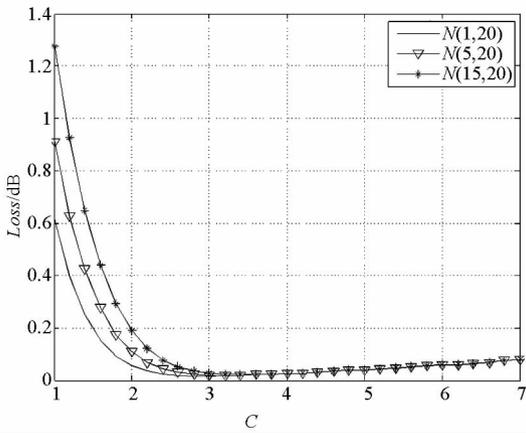


图2 4比特量化时,不同信号对应的限幅门限系数与量化损耗的关系

Fig. 2 Three signal's quantization distortion vs. clip level under the same quantization

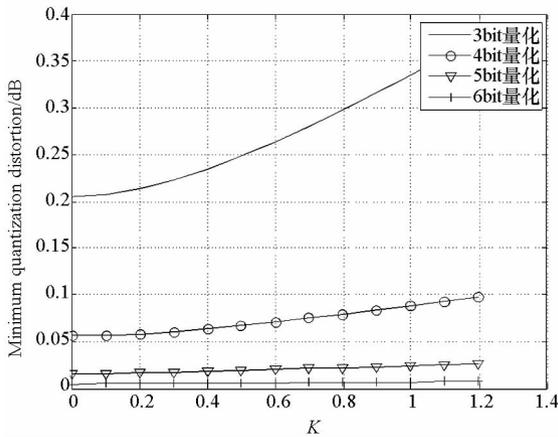


图3 不同量化位数下,信噪比系数对应的最小量化损耗

Fig. 3 SNR coefficient vs. minimum quantization distortion under different quantization

表1 信噪比系数与最佳门限系数关系

Tab. 1 SNR coefficient vs optimum clip level coefficient

信噪比系数 K	门限系数 C_{opt}	最小量化损耗 $Loss_min (dB)$	仿真的最小量化损耗 $Loss_min (dB)$
0	3.27	0.0046	0.0035
0.3	3.4	0.0050	0.0046
0.6	3.65	0.0057	0.0051
0.9	3.93	0.0066	0.0058
1.0	4.03	0.0069	0.0061

以下讨论在量化信号固定时,量化位数与最佳门限系数的关系,对于均匀量化,在量化信号信噪比系数 K 和量化位数 B 固定的情况下,若要达到量化损耗最小,对应的量化最佳限幅门限 A_{opt} 与量化信号幅度均方根 σ 的比值 (C_{opt}) 是一个固定的关系,表2列出了 $K=0.1$ 时,量化位数与最佳门限系数之间的对应关系表,在低比特量化时,

如从3比特量化变成4比特量化,其对应的最小损耗会有0.15dB左右的改善,当量化位数大于7时,其对应的最小量化损耗差别已经不大。

表2 量化位数与最佳门限系数关系

Tab. 2 Quantization bits vs optimum clip level coefficient

量化位数 $B+1$	门限系数 C_{opt}	最小量化损耗 $Loss_min (dB)$	仿真的最小量化损耗 $Loss_min (dB)$
3	1.98	0.2078	0.203
4	2.5	0.0567	0.059
5	2.92	0.0162	0.0225
6	3.29	0.0047	0.0058
7	3.63	0.0014	0.0012
8	3.94	3.87e-4	3.45e-4
9	4.24	1.10e-4	1.03e-4

3 仿真及实验结果分析

为了验证以上分析的正确性,本文采用蒙特卡洛仿真,生成高斯分布的一组10000个随机信号,对信号进行量化后比较实际数据的性能特性与理论分析之间的差异。

图4是服从 $N(0,20)$ 、 $N(10,20)$ 、 $N(15,20)$ 分布的信号经过4bit量化后,量化损耗与限幅门限的关系,由图可知在低比特量化时信号的信噪比系数对最小量化损耗是有影响的,当均值由0变化到15时,带来的最小损耗的变化值,理论分析是0.021dB,仿真结果是0.027dB,仿真值与理论分析值重合性很好。图5是相同的信号经过不同量化位数的量化后,限幅门限与量化损耗的关系,从图上可以看出,随着量化位数的增加,量化带来的损耗会减小,其最优限幅门限的取值随着量化位数的增加相应增加。

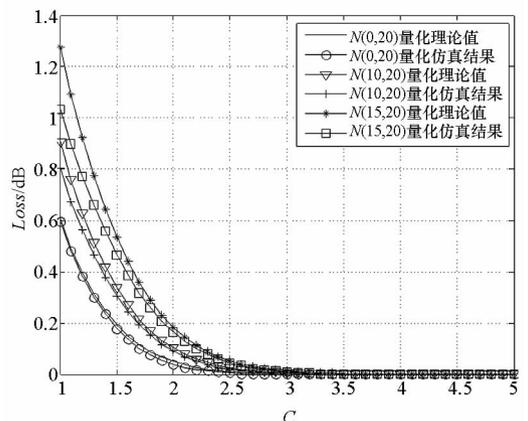


图4 不同信号量化损耗与限幅门限的关系

Fig. 4 Quantization distortion vs clip level under different signals

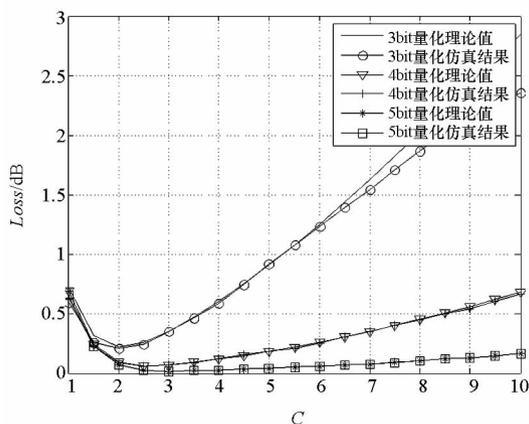


图5 相同信号量化后损耗与限幅门限的关系
Fig.5 Quantization distortion vs clip level under the same signal

同时为了对文中的方法进行进一步验证,使用 Max^[8]提出的方法计算出 $N(0,1)$ 标准高斯分布信号的最优量化步长与最小均方误差,与本文解析方法计算的结果进行对比,如表3所示。从表中的结果可见,本文的方法与 Max 的方法计算结果几乎重合。

表3 本文计算结果与 Max 的计算结果比较

Tab.3 Comparison of mean-square error and output spacing for two methods

量化位数 $B + 1$	Max 方法 的最小 均方差	本文的 最小均 方差	Max 方法 的最优 量化步长	本文的 最优量 化步长
3	0.0375	0.0416	0.586	0.612
4	0.0116	0.0127	0.335	0.352
5	0.0035	0.0037	0.188	0.193
6	0.0011	0.0011	0.104	0.105
7	$3.0467e - 4$	$3.0864e - 4$	0.057	0.057
8	$8.8684e - 5$	$8.8312e - 5$	0.031	0.031
9	$2.5279e - 5$	$2.5008e - 5$	0.017	0.017

4 结论

本文使用概率分布方法,从理论上推导出了 $N(\mu, \sigma^2)$ 高斯随机信号量化门限的取值与量化损耗的关系,以上推导方法同样适用于其他分布的随机信号。通过分析,可以得出 $N(\mu, \sigma^2)$ 高斯分布的随机信号在 AD 量化时量化损耗与限幅门限的一些对实际工程具有指导意义的结论:

(1) 限幅门限的取值对信号量化后的损耗有影响,存在一个最优门限,使得量化后信号的损耗最小。限幅门限通常在工程中的经验取值 3σ , 在低比特量化时,实际会带来较大的偏差,如 3bit 量化,经验门限值与最优门限值对应的损耗有 0.15dB 的差异,但在高比特量化时(大于 5 比特),经验门限值与最优门限值对应损耗的差别不大,通常小于 0.001dB。

(2) 信号的信噪比系数 $K(K = \mu/\sigma)$ 对量化的最小损耗有影响,信噪比系数 K 越大,其对应的最小损耗也越大,在低比特量化时,如 3 比特量化,当 K 由 0 增大至 1 时,带来的损耗增加了 0.13dB;但随着量化位数的增加,其对损耗的影响将会减小,在 6 比特量化时, K 由 0 增大至 1,带来的损耗只增加了 0.002dB 左右,可以忽略。

参考文献:

- [1] 王世练,张尔扬.直扩数字接收机中 AD 量化比特数的确定[J].通信学报,2004,25(8):124-128.
- [2] Shen B, Zhang Q L. A New Method for Analyzing the Quantization Effect of ADC in Broadband QAM Receiver[J]. International Conference on Communications, Circuits and Systems, 2002: 1262-1266.
- [3] Guan Z Y, Hulyalkar S N. Bit-Precision Requirements on the A/D Converter in a QAM Receiver[J]. IEEE Transaction of Consumer Electronics, 39(3), 1993: 692-695.
- [4] Chang H. Presampling Filtering, Sampling and Quantization Effects on the Digital Matched Filter Performance [J]. Proceedings of International Telemetering Conference, 1982: 889-915.
- [5] Oppenheim A V, Schaffer R W, Buck J R. Discrete-time Signal Processing[M]. 2nd Ed. Prentice-Hall, Inc. 1999: 157-160.
- [6] 郭南,洪福明,等.软限幅效应,量化阶数及取样间隔对直扩数字匹配滤波性能的影响[J].通信学报,1996,17(1):12-17.
- [7] Lu Y E, Yan Y G, Yuan H P, et al. The Effect of the ADC Quantization on the Performances of GPS Receiver Adaptive Antenna[J]. IEEE Aerospace Conference, 2007: 1-6.
- [8] Max J. Quantizing for Minimum Distortion [J]. IRE Transactions on Information Theory, 1960: 7-12.
- [9] Sripad A, Snyder D. A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White[J]. Acoustics Speech and Signal Processing, 1977, 25(5):442-448.