УДК 681.3

АППРОКСИМАЦИЯ ФУНКЦИИ СКОРОСТЬ-ИСКАЖЕНИЕ ПРИ ОЦЕНКЕ БИТОВЫХ ЗАТРАТ В СИСТЕМАХ СЖАТИЯ АНАЛОГОВОЙ ИНФОРМАЦИИ

Поров А.В.

Решается задача нахождения функции скорость-искажение для скалярного квантования случайных величин, распределенных по обобщенному гауссову закону. Предлагается способ оценки функции скорость-искажение на основе полиномиальной аппроксимации в области низких скоростей кодирования и использования точного значения для области высоких скоростей при заданной ошибке квантования и параметров распределения случайной величины.

Введение

Оценка скорости кодирования по заданным параметрам весьма актуальна в связи с применением в разнообразных алгоритмах сжатия, в частности, при определении частотно-временной разрешающей способности, оптимальном распределении битовых ресурсов по спектральным диапазонам, использовании переменной битовой скорости кодирования и в некоторых других блоках кодирования аудио информации [1]. Интерес к обобщенному гауссовскому распределению вероятностей обусловлен тем, что многочисленные экспериментальные исследования убеждают в применимости распределений этого класса для описания моделей случайных величин в задачах кодирования аудио и видео информации. Соотношение между ошибкой и скоростью кодирования рассмотренного ниже способа ближе к реальным соотношениям, чем приведенные в стандартах MPEG1,2]. Подробный обзор результатов в области поиска функции скорость-искажение скалярного квантования приведен в работах [3,4,5].

Приведем необходимые определения и обозначения. Пусть $f(x)$ есть плотность распределения вещественной случайной величины x. *Квантователь* задает отображение $y = Q(x)$ множества значений случайной величины x на дискретное (конечное или счетное) множество *аппроксимирующих значений* $Y = \{y\}$. Если заданы распределение $f(x)$ и квантование $Q(x)$, то определены вероятности

$$p(y) = \int\limits_{x:Q(x)=y} f(x)dx,$$

энтропия

$$H(Y) = -\sum_y p(y)\log p(y)$$

и среднеквадратическое искажение

$$D(Q) = \mathrm{E}[(x - Q(x))^2].$$

Из теории кодирования источников без потерь известно, что при использовании кодов переменной длины скорость кода источника с вероятностями букв $\{p(y)\}$ может быть сделана сколь угодно близкой к $H(Y)$. Функция

$$R_s(D) = \inf_{Q:D(Q)\le D} H(Y)$$

представляет собой нижнюю границу скорости кода для случайной величины x при заданном искажении D и называется *функцией скорость-искажение при скалярном квантовании.*

Задача данной работы – вычислить оценку $\widehat{R}_s(D)$ для рассматриваемого класса источников как можно более близкую к $R_s(D)$.

Нижней границей для $R_s(D)$ является теоретико-информационный предел, который определен для последовательности независимых одинаково распределенных величин:

$$H(D) = \min_{f(y|x):\mathrm{E}\left[(x-y)^2\right]\le D} I(X;Y),$$

где $I(X;Y)$ обозначает среднюю взаимную информацию между $X = \{x\}$ и $Y = \{y\}$ при заданном совместном распределении вероятностей $f(x,y) = f(x)f(y\,|\,x)$.

Предложенные ранее способы оценки функции скорости-искажения [1,2] допускают целый ряд достаточно грубых приближений: не учитывают изменения во времени распределения кодируемых данных, предполагают использование только равномерного скалярного квантования и используют весьма приблизительную оценку $\widetilde{R}_s$ скорости при ошибке D:

$$\widetilde{R}_s(D) = H_0(X) - \frac{1}{2}\log(12D), \tag{1}$$

где $H_0(X)$ относительная энтропия выбранной модели распределения, вычисляемая по формуле:

$$H_0(X) = -\int\limits_{-\infty}^{+\infty} f(x)\log_2 f(x)dx.$$

Погрешность оценки функции скорость-искажение по формуле (1) может составлять до 12 дБ, если способ квантования и модель распределения данных не учитываются, и до 11 дБ, если не учитывается только способ квантования.

В данной работе рассматривается еще один подход к построению оценки функции скорость-искажение скалярного квантования. Его важным преимуществом является то, что он описывается не только требуемой ошибкой кодирования, но и параметрами выбранной модели распределения. Не менее важным преимуществом является использование известных теоретических результатов только для области высоких скоростей и полиномиальной аппроксимации скорости для области низких скоростей кодирования, где теоретические оценки чрезвычайно не точны. Погрешность предлагаемого способа оценки функции скорость-искажение не превышает 1 дБ во всем диапазоне скоростей

Модель источника и границы функции скорость-искажение скалярного квантования

Для дальнейшего изложения материала необходимо выбрать модель источника данных. Пусть источник порождает случайную величину, плотность распределения вероятностей которой определена обобщенным гауссовым распределением:

$$f(x)=\frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\exp\left\{-\left(\eta(\alpha,\sigma)|x-m|\right)^{\alpha}\right\},$$

где m,σ – математическое ожидание и среднеквадратическое отклонение, α – параметр распределения, $\Gamma(\cdot)$ – гамма функция, а значения функции $\eta(\alpha,\sigma)$ вычисляются по формуле

$$\eta(\alpha,\sigma)=\sigma^{-1}\left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{1/2}.$$

Параметр α характеризует экспоненциальную скорость убывания хвостов плотности. Влияние параметра α на функцию плотности вероятности (ФПВ) приведено на рис. 1. Значениям $\alpha=1,2$ соответствуют распределение Лапласа и нормальное распределение.

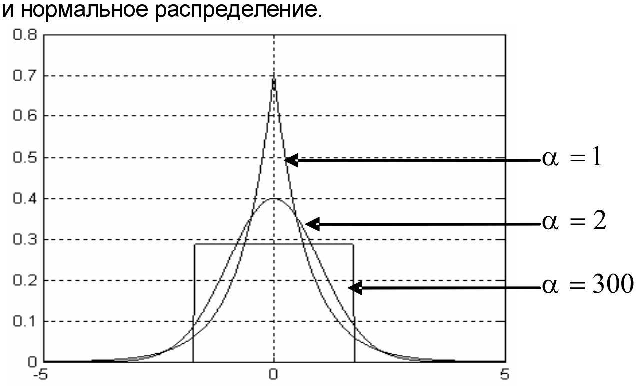


Рис. 1. ФПВ обобщенного гауссовского распределения

Аналитическая форма функции скорость-искажение для большинства распределений вероятностей неизвестна, поэтому в качестве ее нижней границы часто используют границу Шеннона (см.[3, 4])

$$H(D)\ge H_0(X)-\frac{1}{2}\log_2(2\pi eD)=H_{Sh}(D), \qquad ()$$

где D – среднеквадратичная ошибка.

При заданных параметрах α и σ относительная энтропия обобщенного гауссовского распределения выражается аналитически [6]:

$$H_0(X)=-\log_2\left[\frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right]+1/(\alpha\ln 2). \qquad (3)$$

Таким образом, если в формуле (1) учесть изменения распределения данных, то оценка $\tilde{\tilde{R}}_s(D)$ функции скорость-искажение примет вид:

$$\tilde{\tilde{R}}_s(D)=-\log_2\left[\frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right]+1/(\alpha\ln 2)-\frac{1}{2}\log(12D). \qquad (4)$$

Из (5) и (6) следует граница Шеннона для обобщенного гауссовского распределения

$$H_{Sh}(D)=-\log_2\left[\frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right]+1/(\alpha\ln 2)-\frac{1}{2}\log_2(2\pi eD).$$

Граница Шеннона $H_{Sh}(D)$ совпадает с функцией скорость-искажение $H(D)$ для гауссовской случайной величины, т.е. при $\alpha=2$. При других значениях α разность $H(D)-H_{Sh}(D)$ может быть довольно большой, но на высоких скоростях кодирования разница будет стремиться к нулю. Приближенные значения функции $H(D)$ можно получить, используя алгоритм Блейхута [7].

В качестве верхней границы для $R_s(D)$ можно принять важный результат, впервые установленный Кошелевым В.Н. [8] (1963 г.), позже Задором [9] (1966 г.), Гишем и Пирсом [10] (1968 г.). При достаточно большой скорости (при малой ошибке квантования) имеют место неравенства

$$H(D)\le R_s(D)\le H(D)+\frac{1}{2}\log\frac{\pi e}{6}\approx H(D)+0.2546,$$

Из этих неравенств следует, что нижняя граница Шеннона является ориентиром при оценке эффективности скалярного квантования, поскольку всегда выполняются неравенства:

$$H_{Sh}(D)\le H(D)\le R_s(D)\le H(D)+0.2546. \qquad (5)$$

Неравенства (5) могут быть записаны в упрощенном виде для области высоких скоростей кодирования, потому что граница Шеннона будет совпадать с функцией скорость-искажение источника и асимптота Кошелева будет точна:

$$H_{Sh}(D)=H(D)\le R_s(D)=H_{Sh}(D)+0.2546. \qquad (6)$$

Из неравенств (6) следует, что при малых ошибках квантования (высокой скорости кодирования) функция скорость-искажение скалярного квантования определяется нижней границей Шеннона и асимптотой Кошелева:

$$R_s(D)=H_{Sh}(D)+0.2546. \qquad (7)$$

Таким образом, в общем случае функцию скорость-искажение скалярного квантования для малых ошибок можно оценить согласно выражению (7). Для больших ошибок квантования можно применить аппроксимацию как функцию скорости от параметров распределения и ошибки квантования:

$$R_s(D)=\begin{cases}H_{Sh}(D)+0.2546, & D<D_b\\ f(\mathbf{p},D), & D\ge D_b\end{cases}.$$

где D_b - порог по ошибке скалярного квантования, определяющий момент совпадения теоретических результатов с функцией скорость-искажение выбранного способа скалярного квантования, $\mathbf{p}$ - вектор параметров модели распределения данных, $f(\mathbf{p},D)$ - неизвестное аналитическое выражение функции скорость-искажение.

Вектор параметров для обобщенного гауссова распределения содержит параметры α и σ. Скорость и ошибка кодирования не зависит от величины дисперсии, поэтому можно принять $\sigma^2=1$.

$$R_s(D)=\begin{cases}-\log_2\left[\frac{\alpha\eta(\alpha,1)}{2\Gamma(1/\alpha)}\right]+1/(\alpha\ln 2)-\frac{1}{2}\log_2\left(2\pi eD\right)+0.2546, & D<D_b\\ f(\alpha,D), & D\ge D_b\end{cases}. \qquad (8)$$

Функция скорость-искажение скалярного квантования с расширенной нулевой зоной

Формула (8) показывает, что расчет функции скорость-искажение квантования с расширенной нулевой зоной можно производить по двум формулам. Выбор формулы определяется пороговым значением требуемой ошибки D_b. Отметим, что величина порога ошибки D_b зависит исключительно от параметра распределения α. Первая формула из (8) определяет участок кривой, на котором справедливы теоретические результаты в области малых ошибок, вторая – описывает область больших ошибок.

Используя полиномиальную аппроксимацию для нахождения зависимости $f(\alpha, D)$ и оценки порога D_b, оценку $\hat{R}_s(D)$ выражения (8) можно записать следующим образом:

$$\hat{D}_b = b_n\alpha^n + \ldots + b_1\alpha + b_0$$
$$\hat{R}_s(D) = \begin{cases} -\log_2\left[\dfrac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right] + 1/(\alpha\ln 2) - \dfrac{1}{2}\log_2(2\pi e D) + 0.2546, & D < \hat{D}_b \\ c_n(\log_2(D))^n + \ldots + c_1\log_2(D) + c_0 \end{cases} \quad (9)$$

Отметим, что значения коэффициентов c_i также будут зависеть от параметра α, поскольку для каждого распределения данных они уникальны. Коэффициенты аппроксимирующего полинома и величина порога по ошибке D_b для различных параметров распределения α может быть представлена таблицей, но на практике более удобно использовать аналитические зависимости $c_i = c_i(\alpha)$.

Зависимость порога D_b от параметра α является нелинейной функций. Целесообразно использовать разные аппроксимирующие функции для двух диапазонов, поскольку при параметрах $\alpha > 1.5$ функция плотности вероятности чрезвычайно быстро теряет экспоненциальную скорость убывания хвостов плотности. Таким образом, зависимость порога D_b от параметра α становится линейной по мере увеличения значения параметра α. Результаты экспериментов показали, что для параметра $\alpha \le 1.5$ лучше использовать аппроксимирующий полином второй степени, а для других случаев – первой степени (см. рис. 2):

$$D_b = \begin{cases} -0.0406\alpha^2 + 0.1210\alpha - 0.0242, & \alpha \le 1.5 \\ -0.0011\alpha + 0.0683, & \alpha > 1.5 \end{cases} \quad (10)$$

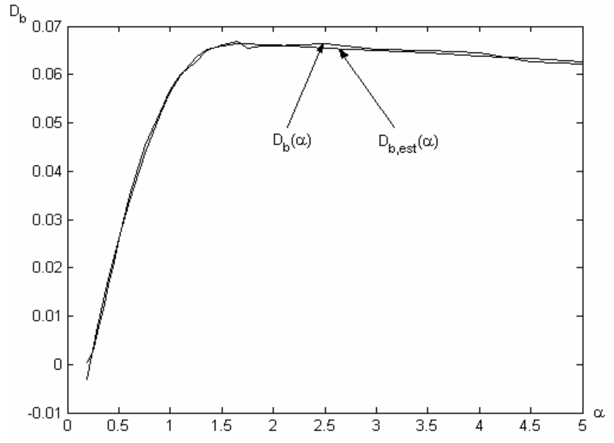


Рис. 2. Зависимость $D_b(\alpha)$ и ее аппроксимация $\hat{D}_b(\alpha)$

Рассмотрим зависимость коэффициентов аппроксимирующего полинома от параметра α. На рис. 3 показаны три зависимости $c_2(\alpha)$, $c_1(\alpha)$ и $c_0(\alpha)$, а также их аппроксимации. Так как при параметрах $\alpha > 1.5$ функция плотности вероятности быстро теряет экспоненциальную скорость убывания хвостов плотности, то можно использовать аппроксимирующий полином первой степени, а для параметра $\alpha \le 1.5$ достаточно использовать полином третьей степени.

$$\hat{c}_2 = \begin{cases} 0.1187\alpha^3 - 0.3266\alpha^2 + 0.2149\alpha - 0.0085, & \alpha \le 1.5 \\ -0.0104\alpha - 0.0293, & \alpha > 1.5 \end{cases}$$
$$\hat{c}_1 = \begin{cases} 0.1088\alpha^3 + 0.0874\alpha^2 - 0.9159\alpha + 0.1464, & \alpha \le 1.5 \\ -0.0195\alpha - 0.7224, & \alpha > 1.5 \end{cases} \quad (11)$$
$$\hat{c}_0 = \begin{cases} -0.0958\alpha^3 + 0.3575\alpha^2 - 0.3472\alpha + 0.0326, & \alpha \le 1.5 \\ 0.0387\alpha - 0.0694, & \alpha > 1.5 \end{cases}$$

Совмещая выражения (9) и (11) для оценки функции скорость-искажение и коэффициентов полинома, с учетом построенной оценки порога по ошибке, получим:

$$\hat{D}_b = b_n\alpha^n + \ldots + b_1\alpha + b_0$$
$$\hat{R}_s(D) = \begin{cases} -\log_2\left[\dfrac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\right] + 1/(\alpha\ln 2) - \dfrac{1}{2}\log_2(2\pi e D) + 0.2546, & D < \hat{D}_b \\ \hat{c}_n(\alpha)(\log_2(D))^n + \ldots + \hat{c}_1(\alpha)\log_2(D) + \hat{c}_0(\alpha) \end{cases}. \quad (12)$$

где $\hat{D}_b$ – вычисляется с помощью формулы (10), $\hat{c}_i(\alpha)$ – вычисляется по формуле (11).

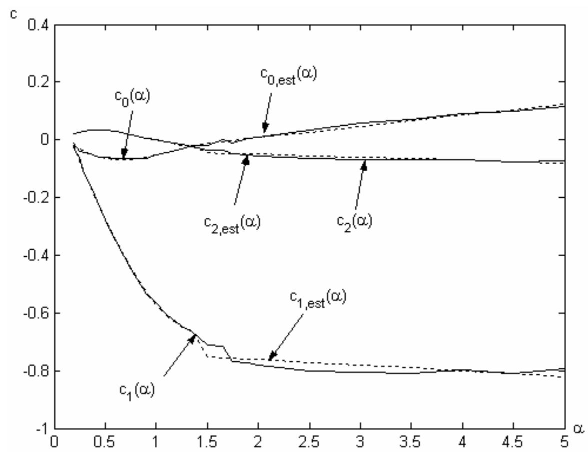


Рис. 3. Зависимости $c_2(\alpha)$, $c_1(\alpha)$ и $c_0(\alpha)$, а также их аппроксимация

Сравнительный анализ полученных результатов

Наиболее часто приводимая в литературе оценка функции скорость-искажение [1,2] основана на нескольких предположениях, некоторые из них могут быть существенными. Важными ограничениями являются предположения, что в кодировании используется равномерное скалярное квантование и неизменное во времени кодирование. В первом стандарте [1] действительно используется равномерное скалярное квантование, и кодирование неизменно во времени, но в последнее время почти все стандарты сжатия мультимедиа информации используют или неравномерное скалярное квантование [2] или векторное [11] и адаптивное кодирование во времени. Выражение (1) позволяет оценить функцию скорость-искажение при условии указанных ограничений. Снимая ограничения, можно существенно улучшить оценку скорости и, как следствие, улучшить алгоритмы, использующие этот критерий. Выражение (12) снимает ограничения на способ квантования и модель источника данных.

Сравнительный анализ оценок функции скорость-искажение проведем на основе вычисления логарифмической погрешности оценки по отношению к оригинальной функции скорость-искажение. Энергетическая эффективность измеряется в децибелах и вычисляется по формуле

$$G = 10\log_{10}\frac{\sigma^2}{D} \text{ (дБ)},$$

где σ^2 – дисперсия источника, а D – ошибка квантования. Без потери общности положим $\sigma^2 = 1$. Тогда максимально достижимый выигрыш $G_{\max}(R)$ при заданной скорости R вычисляется по формуле

$$G_{\max}(R) = -10\log_{10} D_0,$$

где D_0 – корень уравнения

$$R_s(D_0) = R.$$

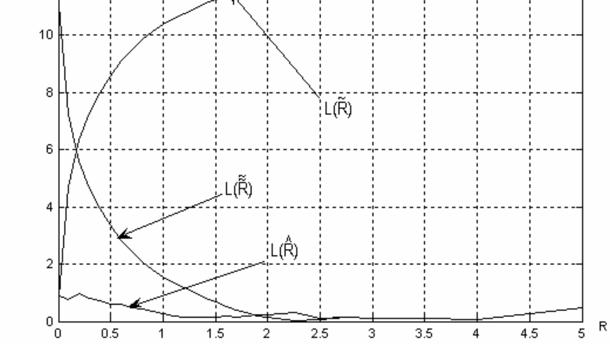
Для оценки скорости $\tilde{R}_s(D)$ и $\hat{R}_s(D)$ можно подсчитать зависимость $D(R)$ как функцию обратную функции скорость-искажение $R(D)$ и вычислить соответствующую энергетическую эффективность:

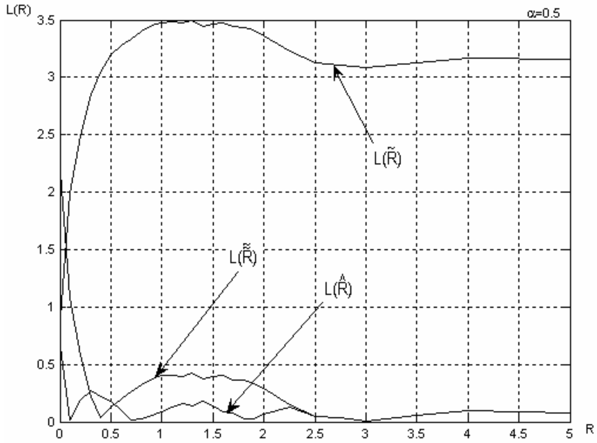
$$G(R) = -10\log_{10} D.$$

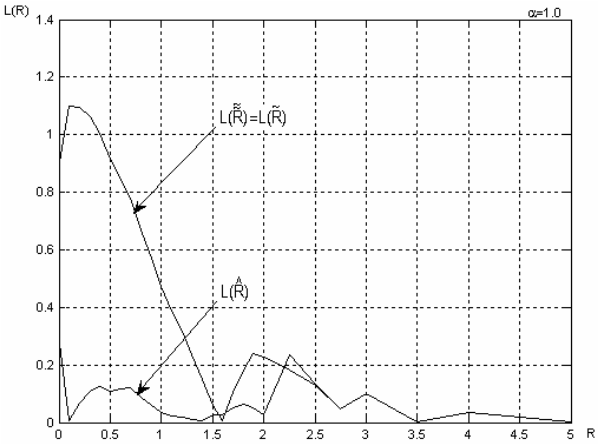
Тогда логарифмическая погрешность рассматриваемой оценки по сравнению с оригинальной функцией скорость-искажение:

$$L(R) = \left|G_{\max}(R) - G(R)\right|.$$

В качестве примера предположим, что моделью является обобщенное гауссово распределение, а способ квантования – квантование с расширенной нулевой зоной [5]. В реальных системах сжатия мультимедиа информации может быть и другая модель и другое квантование, в этом случае необходимо переопределить выражение (12) для заданной модели и способа квантования. Сравнительный анализ оценок функции скорость-искажение по формулам (1) для распределения Лапласа и (12), а также выражений (4) и (12) показывает эффективность уточнения модели и способа квантования при построении оценки функции скорость-искажение. Результаты приведены на рис. 4.







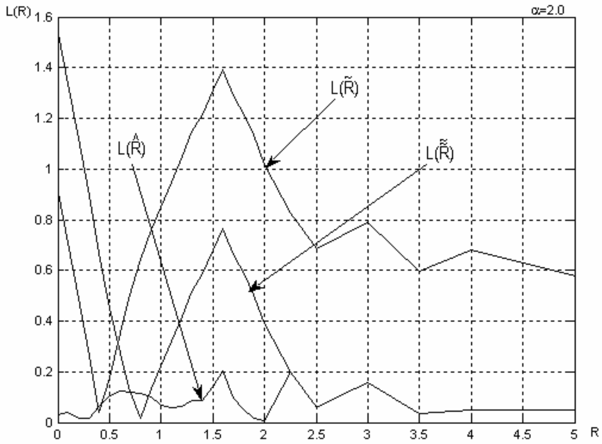


Рис. 4. Логарифмическая погрешность (в дБ) для разных способов оценки функции скорость-искажение скалярного квантования с расширенной нулевой зоной

На рис. 4 логарифмическая погрешность для предложенного способа оценки функции скорость-искажение обозначаются $L(\hat{R})$, для оценки по формуле (4) при определении модели распределения данных обозначаются $L(\tilde{\tilde{R}})$ и для оценки по формуле (1) при неучтенной модели распределения данных обозначаются $L(\tilde{R})$.

Заключение

Предложенный способ оценки функции скорость-искажение скалярного квантования позволяет получить более точную оценку, чем способы, используесые в стандартах MPEG 1 и 2. Сравнительный анализ показал, что существенное уточнение оценки может быть достигнуто при учете модели распределения квантуемых данных, и еще более точная оценка получается, если принять во внимание способ квантования.

При неизвестной модели распределения данных большая логарифмическая погрешность наблюдается во всем диапазоне скоростей, поскольку функция скорость-искажение для различных распределений может существенно отличаться.

Отказ от использования информации о способе квантования при известной модели распределения данных приводит к существенной логарифмической погрешности оценки функции скорость-искажение только в области низких скоростей кодирования, так как для области высоких скоростей (более 3 бит на отсчет) простейшие оценки достаточно точны. Предложенный способ оценки обладает наименьшей логарифмической погрешностью практически во всем диапазоне скоростей, что обусловлено использованием аппроксимирующих функций в области низких скоростей и асимптоты Кошелева в области высоких скоростей кодирования.

Погрешность оценки функции скорость-искажение по известной формуле (1) может составлять до 12 дБ при неизвестном способе квантования и модели распределения данных и до 11 дБ, если не известен только способ квантования. Для предложенного способа оценки функции скорость-искажение погрешность не превышает 1 дБ на всем диапазоне скоростей.

Литература

1. Стандарт IS11172-3 ISO/IEC JTC1/SC29/WG11 MPEG. Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s, Part 3: Audio 1992. (MPEG-1)
2. Стандарт IS13818-3 ISO/IEC JTC1/SC29/WG11 MPEG. Information Technology – Generic Coding of Moving Pictures and Associated Audio, Part 3: Audio 1994. (MPEG-2)
3. R.M. Gray, D.L. Neuhoff. “Quantization,” IEEE Trans. Inform. Theory, vol. IT-44, pp. 2325 –2383, Oct. 1998.
4. Berger T., Gibson J., “Lossy source coding”, IEEE Trans. Inform. Theory, v. 44, No 6, pp. 2702-2703, Oct, 1998.
5. Кудряшов Б.Д., Поров А.В. Скалярные квантователи для случайных величин, имеющих обобщенное гауссовское распределение. Цифровая обработка сигналов. 2005. №4.
6. Sharifi K. and Leon-Garcia A., “Estimation of Shape Parameter for Generalized Gaussian Distributions in Subband Decompositions of Video”, IEEE Trans. on Circuits and Systems for Video Technology, 5(1), pp. 52-56, Febr., 1995.
7. Blahut R. E., “Computation of Channel Capacity and Rate-Distortion Functions”, IEEE Trans. Inform. Theory, 18, No 4, pp. 460-473, Jul., 1972.
8. Кошелев В.Н. Квантование с минимальной энтропией. Проблемы передачи информации. Т.14, стр.151-156, 1963.
9. P. Zador “Topics in the asymptotic quantization of continuous random variables,” IEEE Trans. Inform. Theory, vol. IT-28, pp. 139 – 149, March, 1982.
10. H. Gish, J.N. Pierce, “Asymptotically efficient quantizing”, IEEE Trans. Inform. Theory, v.14, No 5, pp. 676-683, Sept., 1968.