

DOI 10.37539/2949-1991.2026.41.6.036  
УДК 654.191+654.195

**Южанина Алена Александровна**, студентка,  
Сибирский государственный университет  
телекоммуникаций и информатики  
Yuzhanina Alena Aleksandrovna, student,  
Siberian State University of Telecommunications and Informatics

**Шушнов Максим Сергеевич**, к.т.н., доцент,  
Сибирский государственный университет  
телекоммуникаций и информатики  
Shushnov Maxim Sergeevich, PhD, Associate Professor,  
Siberian State University of Telecommunications and Informatics

**О КРИТЕРИЯХ ОЦЕНКИ ВОСПРИЯТИЯ КАЧЕСТВА ЗВУКА  
В РАДИОВЕЩАНИИ НА ОСНОВЕ РЕКОМЕНДАЦИИ ITU-R 1387-2  
ON CRITERIA FOR ASSESSING THE PERCEPTION OF SOUND QUALITY  
IN RADIO BROADCASTING BASED ON RECOMMENDATION ITU-R 1387-2**

**Аннотация.** В статье рассматривается методология оценки воспринимаемого качества звука по стандарту ITU-R BS.1387-2 (PEAQ). Анализируются психоакустические основы метода, версии алгоритма и применение в радиовещании. Особое внимание уделяется критериям интерпретации результатов и их валидации по ITU-R BS.1116-3.

**Abstract.** The article examines the methodology for assessing perceived sound quality according to ITU-R BS.1387-2 (PEAQ). It analyzes the psychoacoustic foundations, algorithm versions, and broadcasting applications. Special attention is paid to interpreting results and their validation according to ITU-R BS.1116-3.

**Ключевые слова:** PEAQ, объективное измерение, ITU-R BS.1387-2, психоакустическое моделирование, радиовещание, ODG.

**Keywords:** PEAQ, objective measurement, ITU-R BS.1387-2, psychoacoustic modeling, broadcasting, ODG.

### 1. Введение

Качество звука является одним из ключевых факторов при проектировании и эксплуатации цифровых систем радиовещания. Лавинообразное внедрение алгоритмов сжатия с низким битрейтом в 1990-х годах поставило перед индустрией сложную задачу: традиционные объективные методы измерений – отношение сигнал/шум (SNR) или коэффициент гармонических искажений (THD) – перестали коррелировать с субъективным восприятием слушателей. Эти методы оказались неприменимы к нелинейным и нестационарным системам, какими являются современные аудиокодеки.

В ответ на это в 1994 году Международный союз электросвязи (ITU-R) инициировал разработку стандартизированного метода объективной оценки воспринимаемого качества. Результатом работы стало принятие в 1998 году (с последующими обновлениями в 2001 и 2023 годах) рекомендации ITU-R BS.1387-2 «Method for objective measurements of perceived audio quality» [1], более известной как PEAQ (Perceptual Evaluation of Audio Quality) [3]. Данная рекомендация предлагает заменить физические измерения физиологически релевантной моделью, предсказывающей оценку, которую дал бы слушатель в контролируемых условиях.



## 2. Обоснование перехода к перцептуальным критериям

Субъективные прослушивания, проводимые по стандарту ITU-R BS.1116-3 [2], являются «золотым стандартом» оценки качества, но они дорогостоящи, длительны и неприменимы для оперативного мониторинга вещания. Метод PEAQ не призван полностью заменить прослушивания, но выступает как эффективный инструмент для:

- непрерывного онлайн-мониторинга тракта вещания;
- оптимизации кодеков при разработке;
- настройки оборудования перед вводом в эксплуатацию;
- скрининга критических тестовых сигналов.

Основная гипотеза PEAQ [2] заключается в том, что качество звука – это не объективная характеристика сигнала, а результат работы сложной биологической системы. Измерения должны имитировать преобразования, происходящие в наружном, среднем и внутреннем ухе (улитке), включая эффекты маскировки (частотной и временной) и нелинейного восприятия громкости.

## 3. Психоакустическая модель PEAQ

PEAQ реализует сквозную модель слуха [1]. На вход подаются два сигнала: *опорный* (исходный, неискаженный) и *тестируемый* (прошедший через измерительное устройство – кодек, тракт). Алгоритм состоит из нескольких этапов. Рассмотрим их составляющие.

### 3.1. Периферическая модель уха

На этом этапе оба сигнала приводятся к «физиологически значимому» представлению: *Взвешивание по кривой Флетчера-Мансона*: Имитируется чувствительность уха к разным частотам (приблизительно от 2 до 5 кГц). Фильтр наружного и среднего уха описывается функцией  $Wd[k]$  [1].

*Частотное шкалирование*: Линейная шкала Гц преобразуется в шкалу Барков (Bark) или ERB (эквивалентная прямоугольная полоса), моделирующую работу базилярной мембраны как банка фильтров. Создается 109 критических полос.

*Частотное и временное расширение (Spreading)*: Моделируется эффект маскировки, когда громкий звук делает неслышимым тихий звук в соседних частотных полосах. Наклоны функции расширения асимметричны (крутой склон в сторону высоких частот).

*Внутренний шум*: Добавляется порог слышимости в тишине – абсолютный порог слуха.

### 3.2. Две версии модели

Рекомендация предлагает два варианта реализации для баланса между точностью и вычислительной сложностью [1]:

*Базовая версия (Basic Version)*: Использует быстрое преобразование Фурье (FFT) с длиной окна 2048 отсчетов (частота 48 кГц) и перекрытием 50%. Разработана для работы в реальном времени (онлайн-мониторинг, п. 3 рекомендации). Использует 11 модельных выходных переменных.

*Расширенная версия (Advanced Version)*: Комбинирует FFT и банк фильтров (анализ вейвлет-типа). Обеспечивает более высокую точность предсказания за счет в ~4 раза больших вычислительных затрат. Применяется для диагностики и анализа при разработке кодеков (пп. 4, 5, 8 рекомендации).

## 4. Система критериев: Модельные выходные переменные (MOV)

PEAQ не выносит вердикт на основе одного параметра. Качество представляет собой многомерную функцию. Модель выделяет из сигнала до 16 различных MOV (Model Output Variables) [1, Attachment 2 to Annex 1], каждая из которых характеризует определенный тип искажения.

### 4.1. Основные группы MOV

1. Анализ модуляций (Bandwidth & Modulation).



Искажения в кодеке часто проявляются не как статический шум, а как «дрожание» громкости во времени.

*BandwidthRefB/BandwidthTestB*: Измеряет срез высоких частот. Потеря высокочастотных компонентов (частота дискретизации) ведет к «глухому» звуку.

*\*WinModDiff1B/RmsModDiffA\**: Оценивают различия в глубине модуляции огибающей сигнала. Критичны для сигналов с атакой (кастаньеты, ксилофон), где возникает эффект «предэха» (pre-echo) — размытие звука во времени перед атакой.

## 2. Шумовые искажения (Noise Loudness)

*RmsNoiseLoudB*: Среднеквадратичное значение громкости внесенного шума (акцент на добавленные компоненты).

*RmsMissingComponentsA*: Характеризует потерю тональных компонентов (частичное маскирование). Используется в асимметричной метрике:

$$*RmsNoiseLoudAsymA = RmsNoiseLoudA + 0.5 \times RmsMissingComponents*$$

*TotNMRB (Total Noise-to-Mask Ratio)*: Отношение шума к порогу маскировки в логарифмическом масштабе. Показывает, насколько заметны ошибки кодирования на фоне маскирующего сигнала.

## 3. Динамические искажения

*AvgLinDistA*: Мера средних линейных искажений (фазовые сдвиги, эхо).

*EHSB (Error Harmonic Structure)*: Анализирует гармоническую структуру ошибки во времени. Помогает выявлять свистящие артефакты (birdies) при низком битрейте.

## 5. Когнитивная модель с переходом от MOV к ODG

Набор MOV имеет разную размерность и физический смысл. Для приведения их к единой шкале субъективного восприятия применяется искусственная нейронная сеть (многослойный перцептрон), обученная на большой базе данных формальных прослушиваний (использовались базы MPEG, ITU, EBU). На выходе нейронной сети получается ODG (Objective Difference Grade), что является количественной мерой, предсказывающей ухудшение качества тестируемого сигнала относительно опорного. Так же определяется соответствие субъективной шкале. Для этого шкала ODG мапшпирруется на шкалу SDG (Subjective Difference Grade) согласно ITU-R BS.1116 [3]. Для этого применяется 5-балльная шкала империпента:

0.0 – неразлично (Imperceptible);

-1.0 – различимо, но не мешает (Perceptible, not annoying);

-2.0 – слегка мешает (Slightly annoying);

-3.0 – мешает (Annoying);

-4.0 – очень мешает (Very annoying).

## 6. Применение критериев в радиовещании

Для целей радиовещания рекомендация выделяет основные классы применения:

Класс 3: Онлайн-мониторинг. Используется *базовая версия*, которая требует использование в реальном времени с минимальной частотой выдачи результатов (ROV)  $\geq 2$  Гц. Это позволяет инженеру видеть «плавающее» качество эфира, предсказывая моменты «провала» кодера на сложных сигналах (например, сибиллянты в речи или удары тарелок).

Класс 1 и 4: Диагностика и разработка. Используется *расширенная версия*, которая позволяет инженерам, анализируя MOV, понять причину низкой оценки: срезаны верхние частоты (*BandwidthTestB*) или идет накопление шума предэха (*RmsNoiseLoudB*).

Класс 8: Поддержка субъективной оценки помогает отобрать «критические» фрагменты (токены) для проведения формальных прослушиваний.

## 7. Ограничения метода

Несмотря на высокую корреляцию с субъективными оценками (до 0.85 для



расширенной версии), PEAQ имеет ограничения, которые инженер вещания обязан учитывать и применять при этом рекомендацию [3]:

*Требование синхронизации:* Сигналы должны быть идеально выровнены во времени. Любая задержка в тракте вызовет ложную тревогу (резкое падение ODG).

*Тестовые сигналы:* Верификация проводилась на конкретном наборе сигналов (Таблица 3 рекомендации [3]: труба, арфа, женская речь и др.). Для синтетических сигналов (например, синусоидальные свипы) предсказания не достоверны.

*Уровень воспроизведения:* Модель чувствительна к уровню громкости (SPL). Для корректной работы требуется калибровка, иначе частотное маскирование будет моделироваться неверно.

## **8. Заключение**

Рекомендация ITU-R BS.1387-2 [1] предоставляет инженерам радиовещания инструмент объективного контроля качества, основанный на физиологии слуха человека. Использование ODG в качестве критерия позволяет перейти от абстрактных «битрейтов» к реальной оценке «слышно-не слышно», что важно для формирования достоверного суждения о реально слышимом/ощущаемом качестве звука на выходе тракта передачи звуковой информации.

Однако автоматизация не отменяет экспертизы. Метод PEAQ эффективен для выявления артефактов (сжатие, предэхо), но менее точен при оценке нелинейных искажений аналоговых трактов или специфических пространственных эффектов в стерео/окружение.

В современной вещательной практике оптимальным является гибридный подход, сочетающий непрерывный объективный мониторинг ODG с последующей выборочной валидацией результатов методом экспертных прослушиваний [3]. Валидация результатов требует применения референсных трактов и соответствующих технических средств и методик проведения исследований, что представляет интересную научную задачу, охватывающую сразу в нескольких смежных научных областях: аудиотехнику, акустику, психоакустику и психофизику восприятия звука.

### *Список литературы:*

1. ITU-R Recommendation BS.1387-2, "Method for objective measurements of perceived audio quality," ITU, Geneva, 2023.
2. Thiede, T., et al., "PEAQ - The ITU Standard for Objective Measurement of Perceived Audio Quality," J. Audio Eng. Soc., vol. 48, pp. 3–29, 2000.
3. ITU-R Recommendation BS.1116-3, "Methods for the subjective assessment of small impairments in audio systems," ITU, Geneva, Geneva, 2015.

