ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 58245— 2018/ IEC/TR 62251:2003

СИСТЕМЫ И ОБОРУДОВАНИЕ МУЛЬТИМЕДИА

Оценка качества. Системы аудио-, видеосвязи

(IEC/TR 62251:2003, IDT)

Издание официальное

Москва
Стандартинформ
2018

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Автономной некоммерческой организацией «Научно-технический центр сертификации электрооборудования «ИСЭП» (АНО «НТЦСЭ «ИСЭП») на основе собственного перевода на русский язык англоязычной версии международного документа, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 452 «Безопасность аудио-, видео-, электронной аппаратуры, оборудования информационных технологий и телекоммуникационного оборудования»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 17 октября 2018 г. № 801-ст

4 Настоящий стандарт идентичен международному документу IEC/TR 62251:2003 «Системы и оборудование мультимедиа. Оценка качества. Системы аудио-, видеосвязи» (IEC/TR 62251:2003 «Multimedia systems and equipment — Quality assessment — Audio-video communication systems», IDT).

Международный документ разработан Техническим комитетом ТС 100 «Аудио-, видео- и мультимедийные системы и оборудование» Международной электротехнической комиссии (IEC).

При применении настоящего стандарта рекомендуется использовать вместо ссылочных международных стандартов соответствующие им национальные стандарты, сведения о которых приведены в дополнительном приложении ДА

5 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Стандартинформ, оформление, 2018

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 2
4 Конфигурация для оценки качества 2
4.1 Входные и выходные каналы 2
4.2 Точки входных и выходных оконечных устройств 3
5 Качество видеосигнала 3
5.1 Введение 3
5.2 Сквозное воспроизведение тонов 4
5.3 Сквозное воспроизведение цвета 6
5.4 Сквозные цветовые разности 7
5.5 Сквозное отношение максимального сигнала к шуму (PSNR) 10
5.6 Сквозная объективная оценка качества видеосигнала 13
6 Качество аудиосигнала 14
6.1 Воспринимаемое качество аудиосигнала по отношению к полным эталонным сигналам 14
6.2 Частота дискретизации и разрешение квантования 17
6.3 Задержка 17
7 Качество видео- и аудиосигналов в целом 18
7.1 Синхронизация аудио- и видеосигнала (синхронизация изображения и речевых сигналов) 18
7.2 Масштабируемость 19
7.3 Общее качество испытуемых сигналов 19
Приложение А (справочное) PSNR, определенные в трехмерных пространствах, применяемые к гипотетическому ухудшению испытуемых сигналов по отношению к исходным эталонным видеосигналам 20
Приложение В (справочное) Сквозная объективная оценка качества видеосигнала в пространственно-частотной области 24
Приложение С (справочное) Описание метода объективного измерения PEAQ 28
Приложение ДА (справочное) Сведения о соответствии ссылочных международных стандартов национальным стандартам 32
Библиография 33

Введение к международному стандарту

Международная электротехническая комиссия (МЭК) — всемирная организация по стандартизации, включающая в себя все национальные комитеты (национальные комитеты МЭК). Цель МЭК заключается в развитии международного сотрудничества по всем вопросам стандартизации в области электрики и электроники. Для этого, кроме осуществления других видов деятельности, МЭК публикует международные стандарты, технические требования, технические отчеты, технические требования открытого доступа (ТТОД) и руководства (далее — публикации МЭК). Их подготовка возлагается на технические комитеты. Любой национальный комитет МЭК, заинтересованный в объекте рассмотрения, может принять участие в этой подготовительной работе. Международные, правительственные и неправительственные организации, сотрудничающие с МЭК, также принимают участие в этой подготовительной работе. МЭК тесно сотрудничает с Международной организацией по стандартизации (ИСО) на условиях, определенных в соглашении между этими двумя организациями.

Официальные решения или соглашения МЭК по техническим вопросам выражают, насколько это возможно, международное согласованное мнение по относящимся к проблеме вопросам, так как каждый технический комитет имеет представителей от всех заинтересованных национальных комитетов МЭК.

Выпускаемые документы имеют форму рекомендаций для международного использования, публикуются в форме стандартов, технических условий, технических отчетов или руководств и принимаются национальными комитетами МЭК именно в таком качестве.

В целях содействия международной унификации национальные комитеты МЭК обязуются максимально ясно и понятно использовать Публикации МЭК в своих национальных и региональных публикациях. Любое расхождение между стандартами МЭК и соответствующими национальными или региональными стандартами должно быть ясно обозначено в последних.

МЭК не предоставляет никакой маркировки соответствия и не несет ответственности за любое оборудование, заявленное как соответствующее одному из ее стандартов.

Необходимо обратить внимание на то, что некоторые элементы данного технического отчета могут быть предметом патентного права. МЭК не несет ответственности за идентификацию частично или полностью такого патентного права.

Основная задача технических комитетов МЭК заключается в подготовке международных стандартов. Тем не менее технический комитет может внести предложение о публикации технического отчета в том случае, когда он собрал данные, отличные от тех, которые обычно публикуются в качестве международного стандарта, например данные, относящиеся к последним техническим достижениям.

Технические комитеты МЭК не обязаны пересматривать технические отчеты до тех пор, пока предоставляемые ими данные не перестанут считаться действительными или полезными для разработчиков документа.

IEC/TR 62251, который является техническим отчетом, был подготовлен Техническим комитетом ТС 100 «Аудио-, видео- и мультимедийные системы и оборудование».

Текст настоящего технического отчета основан на следующих документах:

Проект документа для голосования	Отчет о голосовании
100/561/DTR	100/662/RVC

Полную информацию о голосовании по одобрению настоящего технического отчета можно найти в вышеуказанном отчете о голосовании.

Настоящая публикация была составлена в соответствии с Директивами ИСО/МЭК, часть 2.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

СИСТЕМЫ И ОБОРУДОВАНИЕ МУЛЬТИМЕДИА

Оценка качества. Системы аудио-, видеосвязи

Multimedia systems and equipment. Quality assessment. Audio-video communication systems

Дата введения — 2019—04—01

1 Область применения

Настоящий стандарт устанавливает параметры, которые должны быть измерены объективными методами, а также методы и условия измерений, процедуры обработки измеренных данных и представления полученной информации для объективной сквозной (здесь и далее по тексту термин «сквозной» обозначает «от начала и до конца») оценки качества систем аудио-, видеосвязи, осуществляемой посредством цифровых сетей. Измерения следует проводить при двухстороннем и полном эталонном сигнале. Предполагается, что системы обеспечены каналами электрических интерфейсов на входе и выходе аудио-, видеосигналов для объективной оценки.

Дополнительные требования для систем, не обеспеченных такими каналами, оставлены для дальнейшего изучения.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты. Для датированных ссылок применяют только указанное издание ссылочного стандарта, для недатированных — последнее издание ссылочного стандарта (включая все изменения к нему).

IEC 60268-4, Sound system equipment — Part 4: Microphones (Оборудование звуковых систем. Часть 4. Микрофоны)

IEC 60268-5, Sound system equipment — Part 5: Loudspeakers (Оборудование звуковых систем. Часть 5. Громкоговорители)

IEC 61146-1:1994, Video cameras (PAL/SECAM/NTSC) — Methods of measurement — Part 1: Non-broadcast single-sensor cameras [Видеокамеры (PAL/SECAM/NTSC). Методы измерения. Часть 1. Камеры с одним преобразователем, не предназначенные для телевещания]

IEC 61146-2:1997, Video cameras (PAL/SECAM/NTSC) — Methods of measurement — Part 2: Two- and three-sensor professional cameras [Видеокамеры (PAL/SECAM/NTSC). Методы измерения. Часть 2. Профессиональные камеры с двумя и тремя преобразователями]

IEC 61966-2-1:1999 with Amendment 1 (2003), Multimedia systems and equipment — Colour measurement and management — Part 2-1: Colour management — Default RGB colour space — sRGB [Мультимедийные системы и оборудование. Измерение и управление цветом. Часть 2-1. Управление цветом. Цветовое пространство RGB, используемое по умолчанию — sRGB]

IEC 61966-3:2000, Multimedia systems and equipment — Colour measurement and management — Part 3: Equipment using cathode ray tubes (Мультимедийные системы и оборудование. Измерение и управление цветом. Часть 3. Оборудование, использующее электронно-лучевые трубки)

Издание официальное

IEC 61966-4:2000, Multimedia systems and equipment — Colour measurement and management — Part 4: Equipment using liquid crystal display panels (Мультимедийные системы и оборудование. Измерение и управление цветом. Часть 4. Оборудование, использующее жидкокристаллические дисплейные панели)

IEC 61966-5:2000*, Multimedia systems and equipment — Colour measurement and management — Part 5: Equipment using plasma display panels (Мультимедийные системы и оборудование. Измерение и управление цветом. Часть 5. Оборудование, использующее плазменные дисплейные панели)

IEC 61966-9:2000**, Multimedia systems and equipment — Colour measurement and management — Part 9: Digital cameras (Мультимедийные системы и оборудование. Измерение и управление цветом. Часть 9. Цифровые камеры)

ITU-R BS.1387-1:2001, Method for objective measurements of perceived audio quality (Метод объективных измерений воспринимаемого качества аудиосигнала)

ITU-R BT.601-5:1995, Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios (Параметры студийного кодирования цифрового телевидения для стандартного форматного соотношения 4:3 и широкоэкранного форматного соотношения 16:9)

ITU-T J.144:2001, Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference (Методы объективного измерения воспринимаемого качества видеосигнала для цифрового кабельного телевидения при наличии полного эталонного сигнала)

ITU-T P.931:1998, Multimedia communications delay, synchronization and frame rate measurement (Задержка при передаче мультимедийных данных, синхронизация и измерение частоты кадров)

3 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

3.1 **система аудио-, видеосвязи** (audio-video communication system): Система, обрабатывающая аудио-, видео- и, возможно, другие потоки данных синхронизированным в рамках восприятия пользователей способом для передачи и/или обмена информацией, работающая, как предполагается, через локальную или глобальную цифровую сеть.

3.2 **DMOS** (DMOS): Разность качества исходного и обработанного видеосигнала, оцениваемого на основании средних экспертных оценок (MOS), полученных в результате субъективной оценки испытаний, проведенных группой экспертов по качеству видеосигнала (VQEG).

3.3 **PEAQ** (PEAQ): Оценка восприятия качества аудиосигнала, определенная согласно ITU-R BS.1387-1.

3.4 **PSNR** (PSNR): Объективный параметр качества видеосигнала, определяемый отношением пикового сигнала к шуму, который вычисляют из сравнения исходных и обработанных видеокадров.

3.5 **VQR** (VQR): Объективная оценка качества видеосигнала, заниженная по сравнению с любым объективным параметром благодаря наличию оптимальной корреляции с DMOS.

4 Конфигурация для оценки качества

4.1 Входные и выходные каналы

Аудио- и видеосигнал в аудио-, видеопотоках должны быть зарегистрированы на соответствующих входном и выходном каналах системы аудио-, видеосвязи, как показано на рисунке 1.

---

* Заменен на IEC 61966-5:2008. Однако для однозначного соблюдения требований настоящего стандарта, выраженного в датированной ссылке, рекомендуется использовать только указанное в этой ссылке издание.

** Заменен на IEC 61966-9:2003. Однако для однозначного соблюдения требований настоящего стандарта, выраженного в датированной ссылке, рекомендуется использовать только указанное в этой ссылке издание.

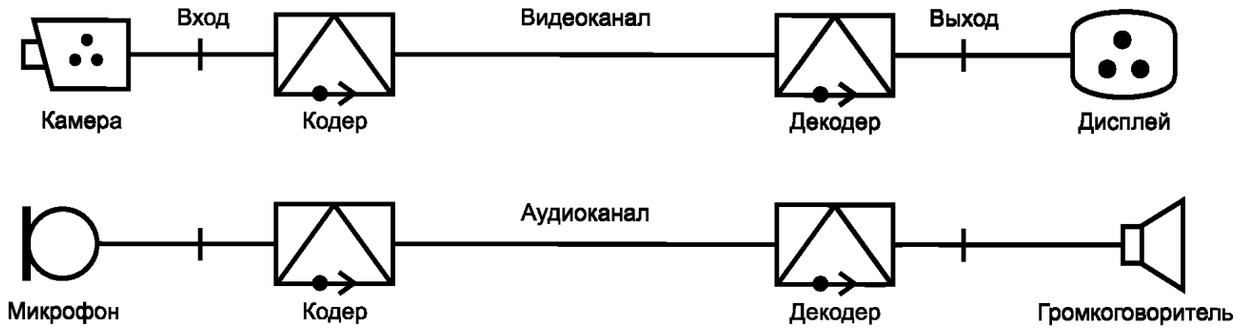
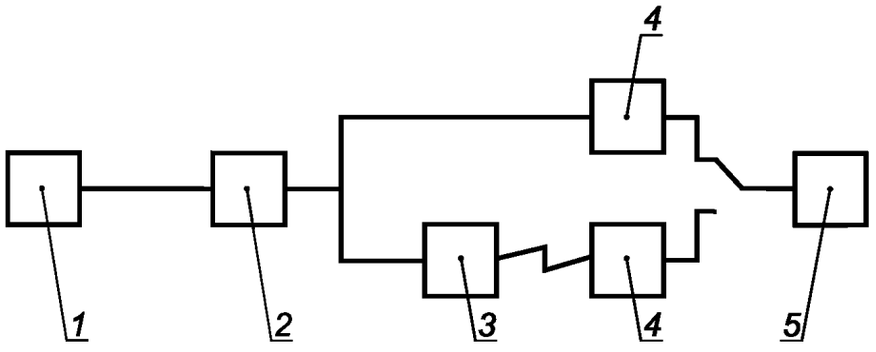


Рисунок 1 — Модель систем аудио-, видеосвязи

4.2 Точки входных и выходных оконечных устройств

Руководствуясь принципом сквозной оценки качества систем аудио-, видеосвязи, точки для получения необработанных данных должны быть в максимально возможной степени предельными оконечными точками. Методы измерения и характеристики оборудования, входящего в состав систем аудио-, видеосвязи, стандартизованы, например на входные преобразователи, такие как видеокамеры и микрофоны, распространяются МЭК 61146-1, МЭК 61146-2, МЭК 61966-9 и МЭК 60268-4, а на выходные преобразователи, такие как дисплеи видеосигнала и громкоговорители, — МЭК 61966-3, МЭК 61966-4, МЭК 61966-5 и МЭК 60268-5, требования которых могут быть не применимы для сквозной оценки качества систем аудио-, видеосвязи.

Структурная схема оценки качества при двухстороннем и полном эталонных сигналах приведена на рисунке 2.



1 — исходный эталонный аудио- или видеосигнал;
2 — предварительная подготовка аудио- или видеосигнала:
- уменьшение динамического частотного диапазона аудиосигнала;
- уменьшение размера кадра и частоты кадров видеосигнала для получения пригодности к оценке качества систем аудио-, видеосвязи, при необходимости;
3 — кодирование с помощью кодера сетевой потоковой передачи с заданной скоростью передачи данных для согласования пропускной способности сквозного сетевого соединения;
4 — декодирование с помощью декодера и визуализация полученных данных для получения слышимых и видимых данных;
5 — сбор данных и расчет оценки качества для предоставления информации, установленной настоящим стандартом

Рисунок 2 — Структурная схема оценки качества

5 Качество видеосигнала

5.1 Введение

В настоящем стандарте для сквозной объективной оценки качества видеосигнала рассмотрены два аспекта:

- статические характеристики, такие как воспроизведение тона и воспроизведение цвета, описанные в 5.2 и 5.3;

- динамические характеристики, основанные на потоковой передаче видеокадров в сети, описанные в 5.4, 5.5 и 5.6.

В качестве исходного эталонного видеосигнала, указанного на рисунке 2 (позиция 1), рекомендуется использовать набор наиболее распространенных исходных видеосигналов, например испытательные последовательности канадского научного центра (CRC). Для фактического кодирования потокового видеосигнала в сеть с ограниченной полосой пропускания исходный эталонный сигнал должен быть уменьшен по размеру кадра и скорости передачи, как указано на рисунке 2 (позиция 2), из-за высокой скорости передачи данных и большого размера кадра, при необходимости.

Доступные в настоящее время эталонные видеопоследовательности для динамических характеристик приведены в таблице А.1. Все источники эталонных видеосигналов, указанные в таблице А.1, которые были использованы группой экспертов по качеству видеосигнала (VQEG) для проведения субъективных испытаний качества видеосигнала в целях получения разности средних экспертных оценок (DMOS), а также для объективной оценки качества видеосигнала (VRG) согласно ITU-R 10-11Q/56-E1, были использованы в настоящем стандарте с разрешения владельца — канадского научного центра (CRC).

Формат каждого из исходных эталонных видеосигналов состоит из 10 кадров (для ракорда) + видеокадры за 8 с + 10 кадров (для трейлера). Существует два видеоформата 525/60 Гц и 625/50 Гц, однако в настоящем стандарте для оценки качества используют только формат 525/60 Гц, приведенный в таблице А.1.

Каждая строка находится в пиксельном мультиплексированном 4:2:2 компонентном видеоформате в последовательности Cb Y Cr Y... и так далее, закодированной в соответствии с ITU-R BT.601-5, где 720 байт/строка для Y, 360 байт/строка для Cb, а Cr — 360 байт/строка для Cr. Строки соединены последовательно в кадры, а кадры соединены последовательно в форме последовательности файлов.

Формат содержит 720 пикселей (1440 байт) на горизонтальную строку и имеет 486 активных строк на кадр. Размеры кадров составляют 1440×486 = 699840 байт/строка, а размеры последовательности — размер файла из 240 кадров за 8 с + 20 кадров. Таким образом, размер файла составляет 699840 байт/кадр × 260 кадров = 181958 400 байт. 30 кадров/с дает скорость передачи данных 699840 байт/кадр × 30 кадров/с × 8 бит = 167961600 бит/с. Поскольку эта скорость передачи данных слишком высокая для обработки обычными персональными компьютерами и потоковой передачи в сети Интернет, исходные испытательные последовательности были уменьшены по размеру кадра до 320×240 пикселей и по формату — до 24-бит/пиксель цветового пространства RGB (вместо цветового пространства YCC) для соответствия типовому видеоформату (AVI) с учетом требований IEC 61966-2-1.

Примечание 1 — Оценка ошибок, проводимая попиксельно, требует уверенного использования крайне высокой степени нормализации. Для нормализации требуется как пространственное, так и временное выравнивание, а также поправки на усиление и смещение. С этой целью следует обратиться к разделу А2 ITU-R 6Q/39-E.

Примечание 2 — Поскольку значения объективных показателей качества в значительной степени зависят от контента видеосигнала, следует в максимально возможной степени использовать различные наиболее распространенные источники видеосигналов.

Примечание 3 — Показатели качества видеосигнала, полученные путем объективной оценки, проведенной согласно разделу 5, должны быть преобразованы в VQR посредством оптимальной корреляции с DMOS, рассматриваемой в ITU-R WP 6Q.

5.2 Сквозное воспроизведение тонов

5.2.1 Параметры, подлежащие оценке

Оценивают сквозную нелинейность в отношении воспроизведения тонов.

5.2.2 Метод оценки

В качестве исходного эталонного сигнала для позиции 1 рисунка 2 должно быть использовано изображение шкалы серых тонов, установленное МЭК 61146-1, как показано на рисунке 3. Статическое ахроматическое изображение должно быть подготовлено (позиция 2 на рисунке 2) и повторно закодировано как потоковый видеосигнал, переданный в сеть.

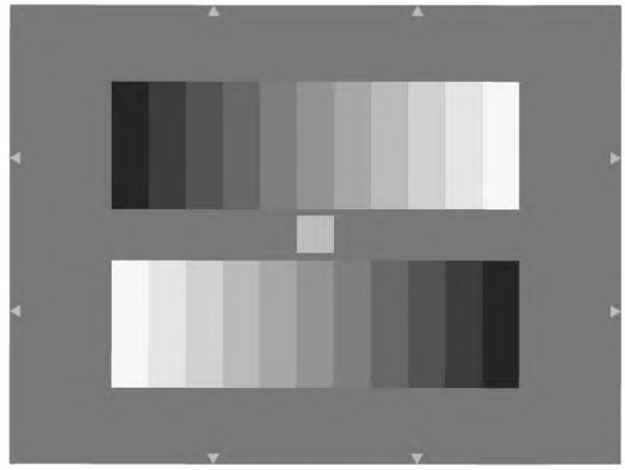


Рисунок 3 — Изображение шкалы серых тонов, установленное МЭК 61146-1

Полученный потоковый видеосигнал должен быть декодирован и воспроизведен устройством просмотра входящих потоковых видеосигналов. Подлежащие отображению данные изображения должны быть зарегистрированы выходным оконечным устройством.

Данные принятого изображения следует сравнивать по показателям трехкомпонентных данных: R (красный), G (зеленый) и B (синий), усредненных в каждой из соответствующих областей.

5.2.3 Представление результата оценки

Результаты оценки, включающие данные воспроизведения изображения по сравнению с входными данными изображения, должны быть представлены в виде таблицы и графика, как показано в таблице 1 и на рисунке 4, соответственно, вместе с оцениваемой системой аудио-, видеосвязи и спецификацией точки входа-выхода.

Таблица 1 — Пример воспроизведения тонов

Измерение	Спецификация			Входной сигнал			Выходной сигнал		
	R, %	G, %	B, %	R	G	B	R	G	B
0	2,0	2,0	2,0	44	43	44	34	39	28
1	4,5	4,5	4,5	63	63	62	55	60	53
2	8,1	8,1	8,1	82	81	82	73	78	69
3	13,0	13,0	13,0	102	102	101	93	98	87
4	19,8	19,8	19,8	123	122	123	115	120	110
5	27,9	27,9	27,9	144	144	144	136	140	128
6	37,8	37,8	37,8	165	164	165	158	163	152
7	48,6	48,6	48,6	184	184	186	174	180	171
8	63,0	63,0	63,0	207	206	208	198	203	195
9	77,3	77,3	77,3	226	227	228	216	219	213
10	89,9	89,9	89,9	243	243	235	217	218	211

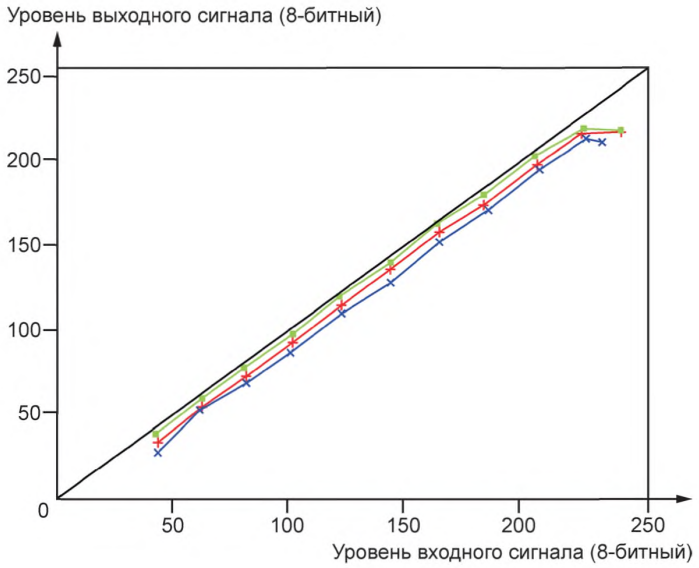


Рисунок 4 — Пример графика воспроизведения тонов

5.3 Сквозное воспроизведение цвета

5.3.1 Параметр, подлежащий оценке

Оценивают сквозные сдвиги цвета в цветовом пространстве CIELAB для статического цветного изображения.

5.3.2 Метод оценки

В качестве исходного эталонного сигнала для позиции 1 рисунка 2 должно быть использовано изображение шкалы воспроизведения цвета, установленное МЭК 61146-1, как показано на рисунке 5. Статическое цветное изображение должно быть подготовлено (позиция 2 на рисунке 2) и повторно закодировано как потоковый видеосигнал, переданный в сеть.



Рисунок 5 — Изображение шкалы воспроизведения цвета по МЭК 61146-1

Полученный потоковый видеосигнал должен быть декодирован и воспроизведен устройством просмотра потоковых видеосигналов. Подлежащие отображению данные цветного изображения должны быть зарегистрированы на выходном оконечном устройстве.

Данные принятого изображения должны быть получены в формате трехкомпонентных данных: R (красный), G (зеленый) и B (синий), усредненных по каждой из соответствующих областей.

5.3.3 Представление результата оценки

Результаты оценки, включающие входные и выходные цвета в формате данных R, G и B, должны быть рассмотрены как цвета в цветовом пространстве sRGB, определенном в МЭК 61966-2-1. Они должны быть преобразованы в CIE 1976 L*a*b* однородное цветовое пространство. Цветовые разности ΔE^*_{ab} между эталонными и полученными данными должны быть рассчитаны и представлены, как показано в таблице 2.

Таблица 2 — Пример воспроизведения цвета

Измерение	Спецификация			Входной сигнал (8-битный×3)			Выходной сигнал (8-битный×3)			Цветовая разность, ΔE^*_{ab}
	R, %	G, %	B, %	R	G	B	R	G	B	
0	87,053	80,546	87,216	222	205	222	221	211	215	3,5
1	48,904	24,181	23,419	184	134	132	186	135	129	1,4
2	37,405	27,352	12,466	163	141	99	164	144	91	4,2
3	25,874	32,782	5,646	138	154	69	139	156	66	2,0
4	12,176	34,717	19,279	98	158	121	96	158	123	0,9
5	15,414	34,081	41,443	109	156	171	109	158	166	2,2
6	17,982	29,222	61,449	117	146	204	119	145	196	1,9
7	36,893	24,007	52,231	164	130	190	163	137	187	3,9
8	51,332	22,896	45,507	188	130	178	187	132	162	5,3
9	43,311	3,062	4,885	174	52	65	172	56	54	8,1
10	83,988	56,759	4,964	236	197	65	219	201	62	4,7
11	2,426	25,943	13,965	47	138	104	45	140	105	1,2
12	3,259	7,178	18,424	54	77	118	50	77	113	2,7
13	82,033	49,052	37,190	233	184	163	219	186	157	3,7
14	10,356	12,908	4,612	91	101	63	89	100	53	5,2
Средняя цветовая разность составляет: $\overline{\Delta E^*_{ab}}$ = 3,396.										

5.4 Сквозные цветовые разности

5.4.1 Параметр, подлежащий оценке

Оценивают среднее значение цветовых разностей в психофизически однородном цветовом пространстве, определенном в МЭС 15.2, между эталонным видеокадром и соответствующим ухудшенным видеокадром.

5.4.2 Метод оценки

В качестве исходного эталонного сигнала для позиции 1 рисунка 2 используют эталонные видеосигналы из таблицы А.1. Видеосигнал должен быть подготовлен путем уменьшения размера кадра видеосигнала в несжатом формате AVI (позиция 2 на рисунке 2). В этой точке необходимо встроить номера кадров, чтобы их можно было использовать для идентификации принятых кадров, соответствующих переданным кадрам.

Закодированные и переданные потоковые видеосигналы должны непрерывно регистрироваться. Вычисления должны быть проведены попиксельно.

Среднюю цветовую разность $\overline{\Delta E^*_{ab_k}}$ в психофизически однородном цветовом пространстве между эталонными и ухудшенными кадрами k рассчитывают по формуле

$$\overline{\Delta E^*_{ab_k}} = \frac{1}{K}\sum_{m=M_1}^{M_2}\sum_{n=N_1}^{N_2}\Delta E^*_{ab_k}, \qquad (1)$$

где $K = \dfrac{1}{(M_2 - M_1 + 1)(N_2 - N_1 + 1)}$ — триплеты в цветовом пространстве CIELAB, соответствующие каждому пикселю эталонного видеокадра k;

$\Delta E^*_{ab_k} = \sqrt{\left(L^*_{d_k} - L^*_{o_k}\right)^2 + \left(a^*_{d_k} - a^*_{o_k}\right)^2 + \left(b^*_{d_k} - b^*_{o_k}\right)^2}$ — цветовая разность CIELAB между пикселями.

Триплеты в цветовом пространстве CIELAB должны быть исключены из значений пикселей *R*, *G* и *B* эталонных и ухудшенных видеокадров по умолчанию в цветовом пространстве RGB (sRGB), определенном МЭК 61966-2-1. Каждый пиксель располагается в строке *m* и столбце *n* видеокадра.

5.4.3 Представление результатов оценки

Результаты оценки представляют в виде графика зависимости цветовой разности между каждым из соответствующих кадров и номерами кадров, как показано на рисунке 6, вместе с идентификацией исходных эталонных видеосигналов. Также должны быть представлены условия измерения, такие как размер кадра в пикселях, частота кадров, скорость потока передачи данных.

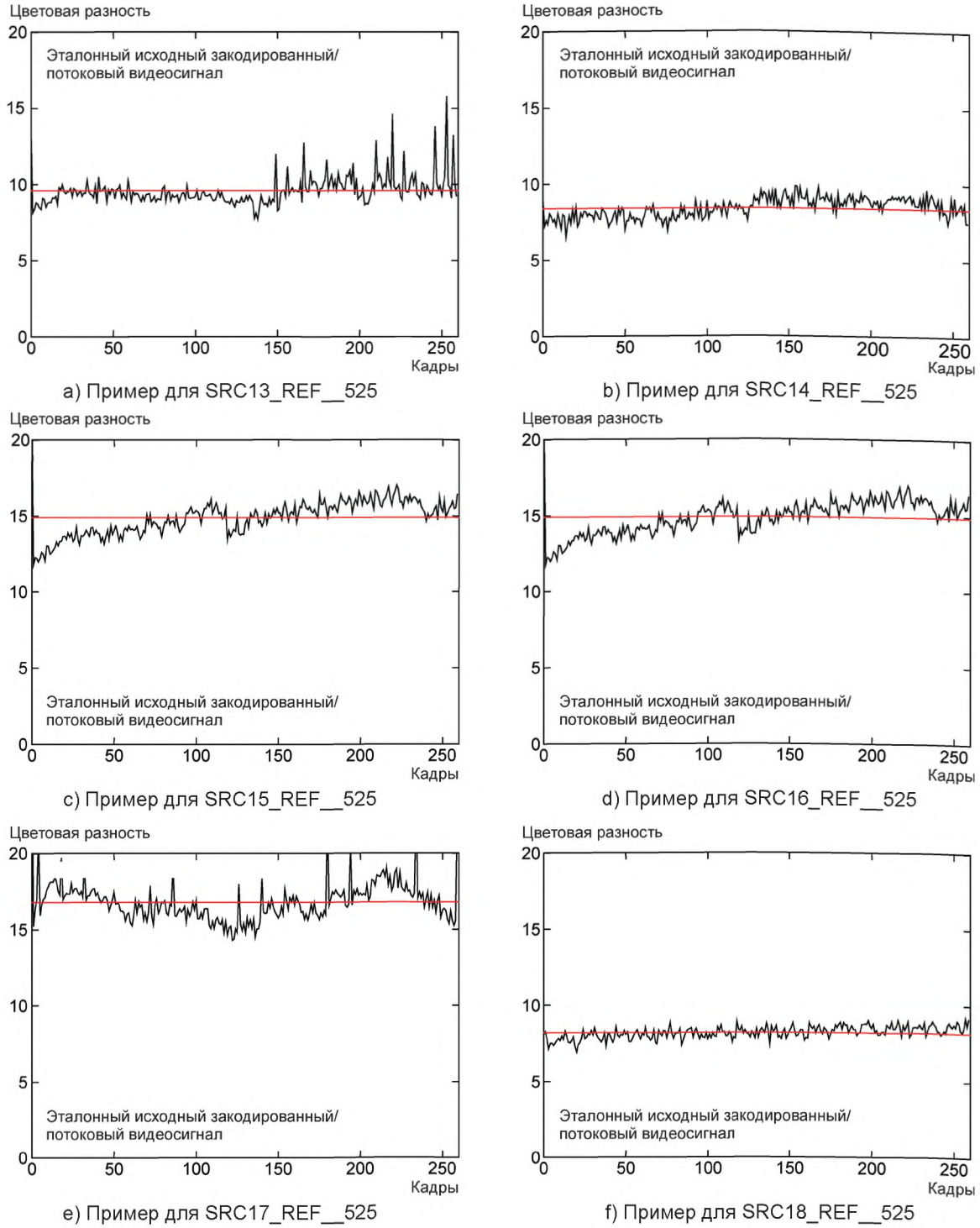
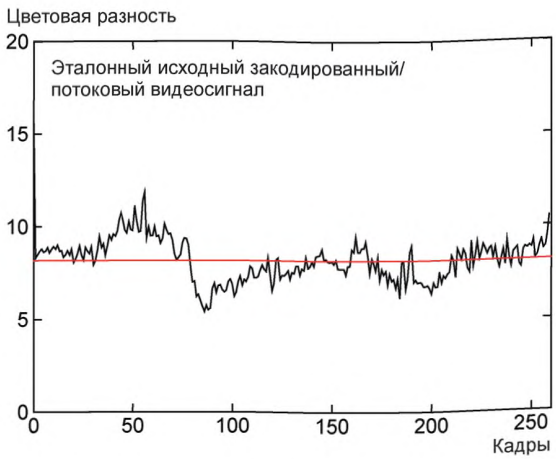
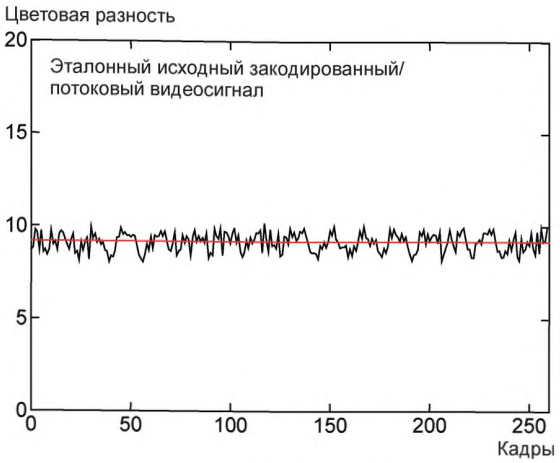


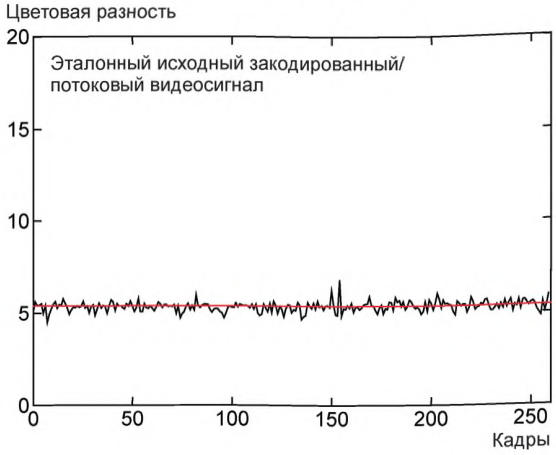
Рисунок 6 — Цветовые разности между эталонными и потоковыми видеокадрами при 250 кбит/с и 30 кадр/с, лист 1



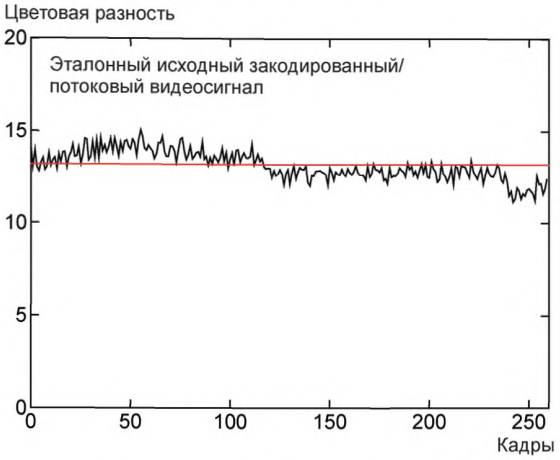
g) Пример для SRC19_REF__525



h) Пример для SRC20_REF__525



i) Пример для SRC21_REF__525



j) Пример для SRC22_REF__525

Условия оценки:
- размер видеокадра — 320 x 240 пикселей;
- частота кадров — 30 кадр/с;
- потоковая скорость передачи данных — 250 кбит/с;
- пропускная способность сети — более 250 кбит/с;
- воспроизведение — плейер Microsoft Media Player ® версия 7.1

Рисунок 6, лист 2

Для заключительной оценки полученные данные должны быть усреднены по кадрам с помощью формулы (2), чтобы обеспечить единичный параметр для объективной оценки, являющийся общим средним значением цветовой разности, которое должно быть представлено, как показано в таблице 3.

$$\overline{\overline{\Delta E}}^{*}_{ab} = \frac{1}{(K_2 - K_1 + 1)} \sum_{k=K_1}^{K_2} \overline{\Delta E}^{*}_{ab_k}, \qquad (2)$$

Таблица 3 — Общие средние значения цветовых разностей

Идентификация исходного эталонного видеосигнала	Общее среднее значение цветовой разности
SRC13_REF__525	9,6
SRC14_REF__525	8,4
SRC15_REF__525	14,9
SRC16_REF__525	8,3

Окончание таблицы 3

Идентификация исходного эталонного видеосигнала	Общее среднее значение цветовой разности
SRC17_REF__525	16,8
SRC18_REF__525	8,2
SRC19_REF__525	8,2
SRC20_REF__525	9,2
SRC21_REF__525	5,4
SRC22_REF__525	13,2

5.5 Сквозное отношение максимального сигнала к шуму (PSNR)

5.5.1 Параметр, подлежащий оценке

Оценивают отношение мощности максимального сигнала к мощности шума, PSNR, в трехмерной системе координат.

5.5.2 Метод оценки

В качестве исходного эталонного сигнала для позиции 1 рисунка 2 используют эталонные видеосигналы из таблицы А.1. Видеосигнал должен быть подготовлен путем уменьшения размера кадра видеосигнала в несжатом формате AVI (позиция 2 на рисунке 2). При необходимости в этой точке встраивают номера кадров, чтобы их можно было использовать для идентификации принятых кадров, соответствующих переданным кадрам.

Закодированные и переданные потоковые видеосигналы должны приниматься непрерывно. Вычисления должны быть проведены попиксельно.

Следует использовать отношение максимального сигнала к шуму (PSNR) между полным эталонным изображением и воспроизведенным изображением, рекомендованное в ITU-T J.144. PSNR рассчитывают по формуле

$$PSNR = 10\lg\left(\frac{S_{max}^2}{MSE}\right), \tag{3}$$

где $MSE = \frac{1}{K}\sum_{p=P_1}^{P_2}\sum_{m=M_1}^{M_2}\sum_{n=N_1}^{N_2}\left(d(p,m,n)-o(p,m,n)\right)^2$,

где $K = \frac{1}{(P_2-P_1+1)(M_2-M_1+1)(N_2-N_1+1)}$;

$d\ (p, m, n)$ и $o\ (p, m, n)$ — представляют, соответственно, ухудшенные и исходные пиксельные векторы в кадре p, строке m и столбце n;

S_{max} — максимально возможное значение пиксельных векторов.

В случае цветных изображений каждый элемент изображения обычно состоит из трехмерных значений: красного (R), зеленого (G) и синего (B). Таким образом, для определения среднеквадратичных ошибок применяют формулу

$$MSE_{RGB} = \frac{1}{K}\sum_{p=P_1}^{P_2}\sum_{m=M_1}^{M_2}\sum_{n=N_1}^{N_2}\left((R_d-R_o)^2+(G_d-G_o)^2+(B_d-B_o)\right)^2, \tag{4}$$

где $S_{max(RGB)} = 3 \cdot 2^{2(N-1)}$ для значений при N-битном кодировании.

Для оценки PSNR в более однородном цветовом пространстве CIE 1976 LAB рекомендуют применять формулу

$$MSE_{Lab} = \frac{1}{K}\sum_{p=P_1}^{P_2}\sum_{m=M_1}^{M_2}\sum_{n=N_1}^{N_2}\left(\Delta E_{ab}^*\right)^2, \tag{5}$$

где $S_{max(Lab)}\Delta E^*_{ab_k} = \sqrt{\left(L^*_{max}\right)^2 + \left(a^*_{max}\right)^2 + \left(b^*_{max}\right)^2}$, фактическое значение которого зависит от цветовой гаммы исходного цветового пространства RGB.

Рекомендуется по умолчанию использовать цветовое пространство RGB, определенное МЭК 61966-2-1, в котором $S_{max(Lab)}$ =148,254.

Примечание — Следует отметить, что члены суммирования в формуле (5) являются квадратом цветовых разностей в психофизически однородном цветовом пространстве, описанных в 5.4.

Дополнительно для сравнения рассчитывают сигнал яркости Y и два цветоразностных сигнала C_b и C_r, обозначенных как Y_{cc}, по формуле

$$MSE_{Y_{cc}} = \frac{1}{K}\sum_{p=P_1}^{P_2}\sum_{m=M_1}^{M_2}\sum_{n=N_1}^{N_2}\left((Y_d - Y_o)^2 + \left(C_{b_d} - C_{b_o}\right)^2 + \left(C_{r_d} - C_{r_o}\right)\right)^2. \quad (6)$$

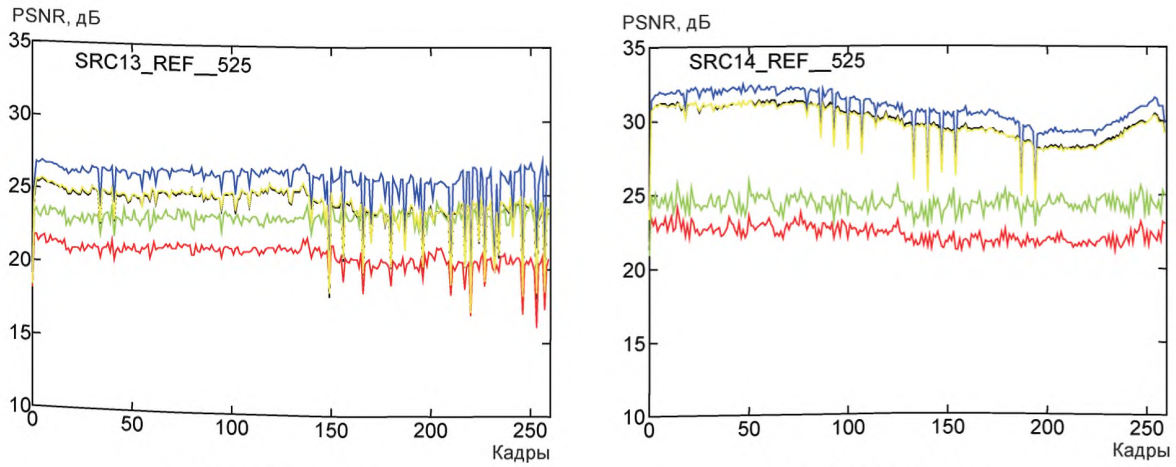
В системе цветового пространства YCbCr, определенного в МЭК 61966-2-1 $S_{max}(Y_{cc})$ = 1,01659.

5.5.3 Представление результатов оценки

PSNR в трехмерных пространствах Lab, Y_{cc} и RGB вместе с PSNR в одномерных пространствах L^* и Y должны быть представлены в протоколе, как показано на рисунке 7.

Также должны быть представлены условия измерения, такие как размер кадра в пикселях, частота кадров, скорость потока передачи данных в битах.

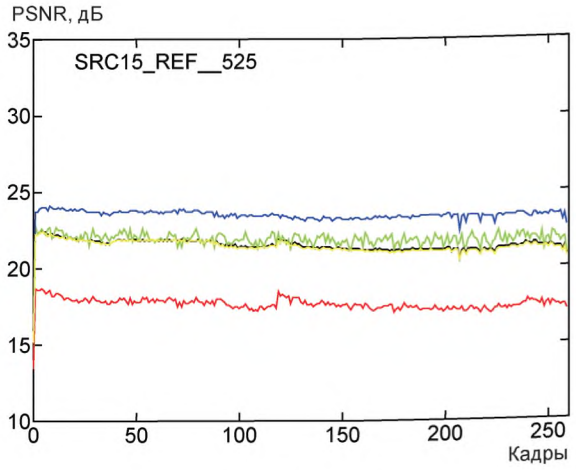
Примечание — В приложении А для информации приведено программное обеспечение для оценки различных параметров качества в отношении известного гипотетического ухудшения, используемого группой экспертов по качеству видеосигнала (GQEG), с точки зрения трехмерных и одномерных PSNR вместе со средней цветовой разностью, разработанное университетом Чиба в сотрудничестве с компанией Mitsubishi Electric Corp.



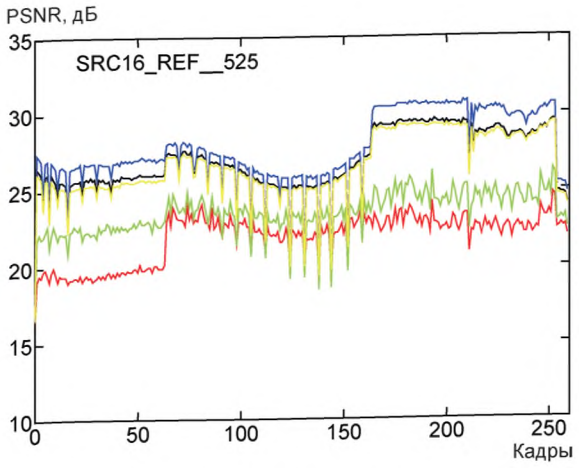
a) SRC13_REF__525

b) SRC14_REF__525

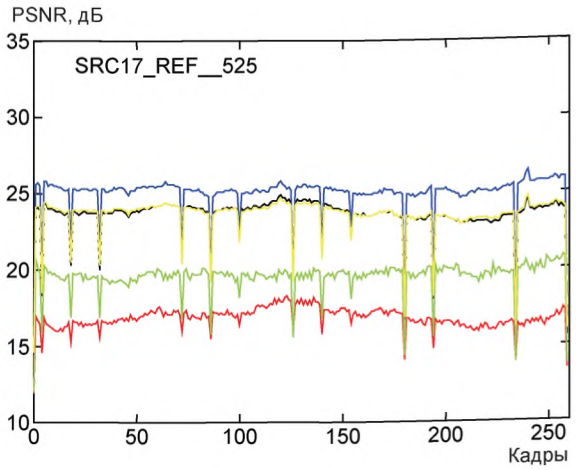
Рисунок 7 — Примеры оценки PSNR, лист 1



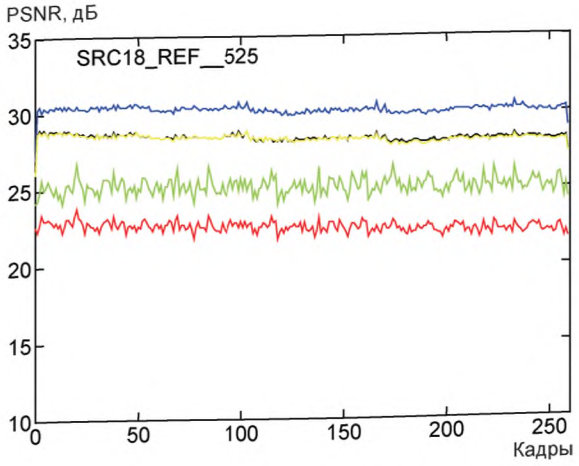
c) SRC15_REF__525



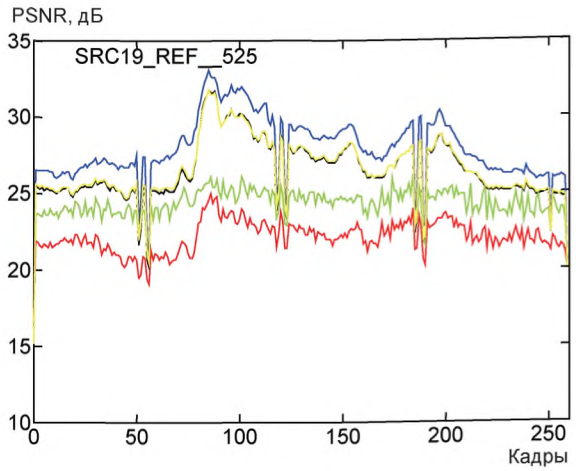
d) SRC16_REF__525



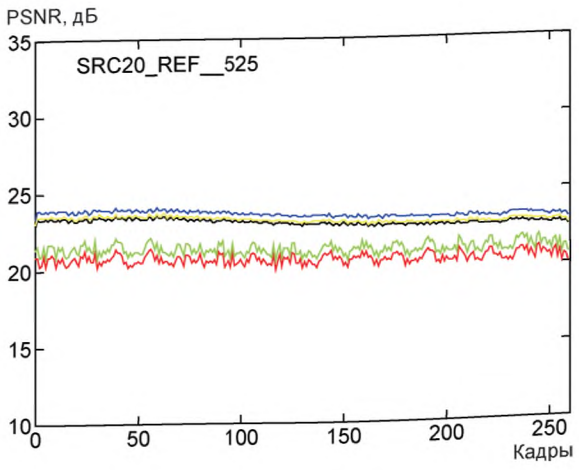
e) SRC17_REF__525



f) SRC18_REF__525

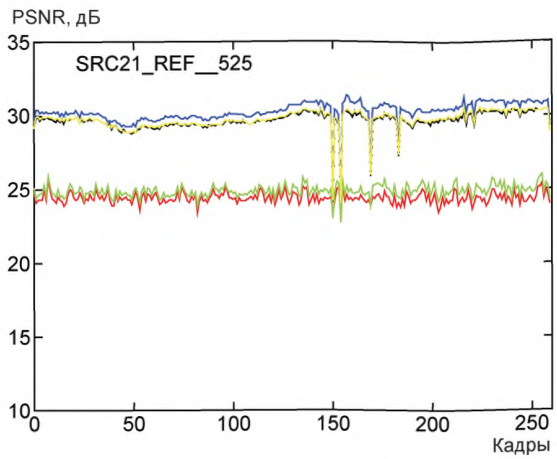


g) SRC19_REF__525

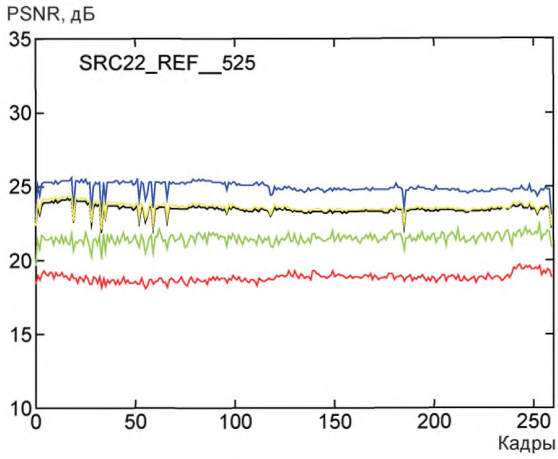


h) SRC20_REF__525

Рисунок 7, лист 2



i) SRC21_REF__525



j) SRC22_REF__525

Условия оценки:
- размер видеокадра — 320 x 240 пикселей;
- частота кадров — 30 кадр/с;
- скорость потока передачи данных — 250 кбит/с;
- пропускная способность сети — более 250 кбит/с;
- воспроизведение — плейер Microsoft Media Player ® версия 7.1

Рисунок 7, лист 3

Для заключительной оценки PSNR должны быть усреднены по кадрам с помощью формулы (7), чтобы обеспечить общие параметры для объективной оценки. Результаты расчета должны быть представлены в виде таблицы 4.

$$\overline{PSNR} = \frac{1}{(K_2 - K_1 + 1)} \sum_{k=K_1}^{K_2} PSNR_K. \quad (7)$$

Таблица 4 — Общие PSNR, усредненные по кадрам

Идентификация эталонного сигнала	PSNR в CIELAB	PSNR в YCC	PSNR в RGB	PSNR в L*	PSNR в Y
SRC13_REF_525	20,9	24,4	24,4	23,3	26,1
SRC14_REF_525	22,3	29,9	30,0	24,4	30,9
SRC15_REF_525	17,7	21,5	21,5	21,9	23,5
SRC16_REF_525	22,1	27,0	27,2	23,7	28,2
SRC17_REF_525	16,9	23,7	23,7	19,6	25,1
SRC18_REF_525	22,5	28,3	28,3	25,2	30,2
SRC19_REF_525	22,3	27,0	27,0	24,4	28,4
SRC20_REF_525	20,7	23,2	23,0	21,3	23,6
SRC21_REF_525	24,4	29,8	29,7	24,9	30,3
SRC22_REF_525	18,8	23,6	23,5	21,5	24,9

5.6 Сквозная объективная оценка качества видеосигнала

5.6.1 Параметр, подлежащий оценке

Оценивают субъективную разность средних экспертных оценок (DMOS) качества видеосигнала с использованием модели, имитирующей характеристики зрения и восприятия человеком цифровых видеосигналов.

5.6.2 Метод оценки

В результате первого этапа испытаний и изучения VQEG в соответствии с ITU-R 10-11Q/56-E предложенных десяти моделей оценки (фактически девять из десяти представленных моделей были признаны эффективными) приняты следующие модели для оценки качества видеосигнала:

a) оценка изображения на основе сегментации, обеспечивающая прогнозирование качества по набору заранее заданных сцен;

b) оценка зрительного различения, имитирующая реакции пространственно-временных зрительных механизмов человека;

c) оценка имитации характеристик зрения человека с использованием пространственно-временных трехмерных фильтров;

d) оценка среднеквадратической ошибки (MSE), взвешенной с помощью зрительных фильтров человека, таких как фильтры на основе пикселей, блоков и последовательностей;

e) оценка параметра искажения восприятия, основанного на пространственно-временной модели зрительной системы человека;

f) оценка, включающая оценку модели восприятия и выделителя признаков, специально настроенного на определенные типы искажений;

g) оценка качества цифрового видеосигнала, включающая множественные аспекты зрительной чувствительности человека при обработке простых изображений;

h) оценка субъективного измерения восприятия видеосигнала с использованием подхода к измерению качества видеосигнала, такого же, как для измерения качества восприятия речи;

i) оценка с использованием характеристик ограниченной полосы пропускания, выделенных из пространственно-временных областей и линейной комбинации параметров для оценки субъективных показателей качества.

Эффективность оценки всех моделей была проверена с точки зрения возможности выделения признаков по сравнению с традиционным методом отношения максимального сигнала к шуму.

VQEG в настоящее время проводит испытания новой предложенной модели для оценки (метода оценки) на основе полного эталонного телевидения. Возможный на практике метод оценки на основе полного эталонного телевидения находится на рассмотрении.

Примечание — Республикой Корея был представлен «новый метод», включающий в себя пространственно-временное вейвлет-преобразование, как описано в ITU-R 6Q/42-E. В настоящем стандарте этот метод был рассмотрен в области цветового пространства sRGB, как показано в приложении В.

5.6.3 Представление результатов оценки

Результаты оценки качества видеосигнала должны быть представлены в виде оценки разности средних экспертных оценок вместе с используемой оцениваемой моделью и условиями.

Примечание — Пример представления результатов оценки находится на рассмотрении.

6 Качество аудиосигнала

6.1 Воспринимаемое качество аудиосигнала по отношению к полным эталонным сигналам

6.1.1 Параметр, подлежащий оценке

Оценивают значения объективной разницы качества (ODG), измеренные методом PEAQ (оценка восприятия качества аудиосигнала), рекомендованным ITU-R BS.1387-1.

6.1.2 Обоснование

Воспринимаемое качество аудиосигнала (PEAQ) является одним из ключевых факторов при разработке цифровых систем аудио-, видеосвязи. Официально принятые испытания на прослушивание всегда были важным методом оценки качества аудиосигнала. Однако субъективные оценки качества являются трудоемкими и дорогостоящими. В связи с этим возникла необходимость разработки объективного метода измерения при проведении оценки качества аудиосигнала. Традиционные объективные методы измерений, такие как отношение сигнала к шуму (SNR) или общее гармоническое искажение (THD), в действительности никогда не показывали достоверного воспринимаемого качества аудиосигнала. Проблемы возникают, когда эти методы применяют по отношению к современным кодекам, которые являются нелинейными и неустановившимися. МСЭ-Р (сектор радиосвязи Международного союза электросвязи) рекомендовал объективный метод измерения для оценки воспринимаемого качества аудиосигнала испытуемого оборудования, например, кодека с низкой скоростью передачи данных, известный как PEAQ (оценка восприятия качества аудиосигнала), Этот метод установлен в ITU-R BS.1387-1 и кратко описан в приложении В.

Выходная переменная объективного метода измерений PEAQ является показателем объективной разницы качества (ODG) и показателем искажения (DI). ODG соответствует показателю субъективной разницы качества (SDG) в субъективной области. Точность оценки ODG ограничена одним десятичным знаком. Однако следует соблюдать осторожность и, в целом, не ожидать, что разница между любой парой ODG в одну десятую показателя будет являться значительной. DI имеет то же значение, что и ODG. Однако DI и ODG можно сравнивать только количественно, но не качественно. Как правило, ODG следует использовать в качестве показателя качества для значений ODG, приблизительно более минус 3,6. ODG очень хорошо коррелируется с субъективной оценкой в этом диапазоне. Если значение ODG составляет менее минус 3,6, следует использовать DI. Таким образом, должны быть измерены обе выходные переменные ODG и DI.

6.1.3 Метод оценки и алгоритм PEAQ

Базовая модель метода объективного измерения PEAQ показана на рисунке 8. Базовая модель состоит из двух входов, один из которых предназначен для (необработанного) эталонного аудиосигнала, соответствующего позиции 2 рисунка 2, а другой — для испытуемого аудиосигнала. Испытуемый аудиосигнал может быть, например, выходным сигналом цифровых систем аудио-, видеосвязи, соответствующим выходу позиции 4 рисунка 2, на который подается эталонный сигнал.

Настоящий метод измерения применим к большинству типов оборудования для обработки цифровых и аналоговых аудиосигналов. В настоящем стандарте рассматривают применение метода только по отношению к цифровым каналам аудиосвязи. Блок «испытуемое устройство» соответствует позициям 2 и 3 рисунка 2.

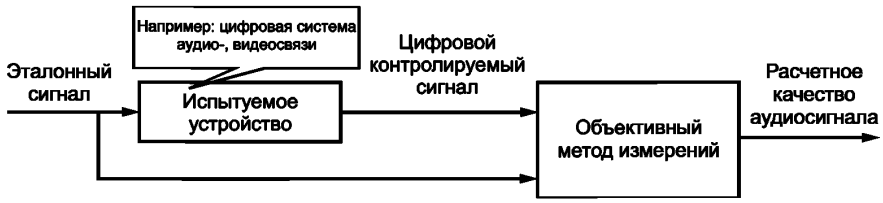


Рисунок 8 — Базовая модель проведения объективных измерений

Алгоритм проведения оценки методом PEAQ приведен на рисунке 9. Метод PEAQ основан на общепринятых психоакустических принципах. В целом он сравнивает сигнал, который был некоторым образом обработан, с соответствующим синхронизированным во времени эталонным сигналом. На первом этапе обработки сигнала используется периферийное устройство, моделирующее ухо, известное как «модель восприятия», или «модель уха». Одновременные блоки данных эталонного и обработанного сигнала преобразуются в выходные сигналы моделей уха. На следующем этапе модели алгоритма определяют присутствие слышимого искажения в испытуемом сигнале путем сравнения выходных сигналов моделей уха. Информация, полученная посредством этих процессов, дает несколько значений, так называемых MOV (выходных переменных метода измерений), и может использоваться для детального анализа сигнала.

Конечной целью является управление показателем качества, состоящего из одного числа, указывающего на слышимость искажений, присутствующих в испытуемом сигнале. В целях сохранения этого параметра требуется некоторая дальнейшая обработка MOV, имитирующая когнитивную часть слуховой системы человека. Поэтому алгоритм PEAQ включает искусственную нейронную сеть.

Существуют две версии PEAQ: «базовая» версия, отличающаяся низким уровнем сложности, и «расширенная» версия, обеспечивающая большую точность за счет более высокой сложности. Структура обеих версий очень похожа и точно соответствует модели PEAQ, показанной на рисунке 9. Основное различие между базовой и расширенной версиями заключается в соответствующих моделях уха и наборе используемых MOV. В приложении С приводится дополнительная информация о PEAQ, которая помогает понять результаты измерений.

Рекомендуется использовать доступные эталонные сигналы из ITU в виде WAV-файлов (в формате Microsoft RIFF) на CD-ROM-диске. Все эталонные сигналы были отобраны на частоте 48 кГц для 16-битного PCM. Эталонные и испытуемые сигналы, предоставленные ITU, уже согласованы друг с другом по времени и уровню, поэтому дополнительного усиления или задержки не требуют.

Алгоритм измерения должен быть откорректирован на уровень прослушивания 92 дБ SPL.

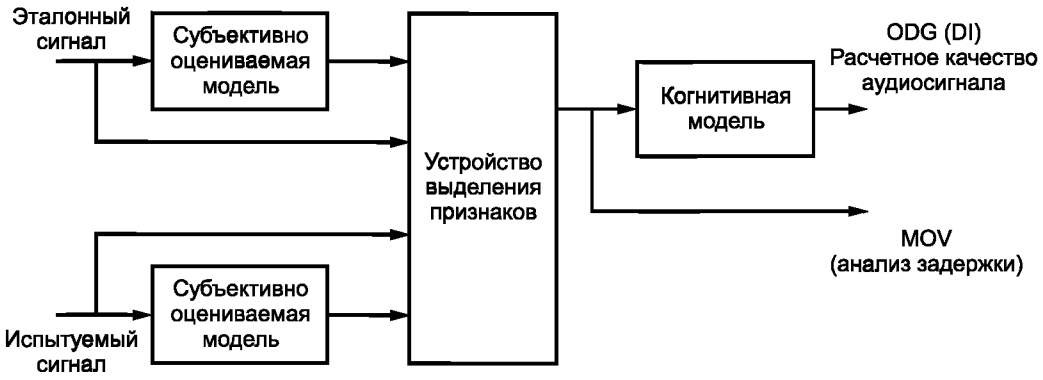


Рисунок 9 — Представление модели PEAQ

6.1.4 Представление результатов оценки

Результаты измерения PEAQ должны быть приведены в виде таблицы параметров указанных эталонного и испытуемого сигналов[1)], а также полученных значений DI и ODG.

Таблица 5 относится к базовой версии, а таблица 6 содержит значения для расширенной версии.

Т а б л и ц а 5 — Параметры испытуемого сигнала и полученные значения DI и ODG для базовой версии

Параметр	DI	ODG
Acodsna.wav	1,304	–0,676
Bcodtri.wav	1,949	–0,304
Ccodsax.wav	0,048	–1,829
Dcodryc.wav	1,648	–0,458
Ecodsmg.wav	1,731	–0,412
Fcodsb1.wav	0,677	–1,195
Fcodtr1.wav	1,419	–0,598
Fcodtr2.wav	–0,045	–1,927
fcodtr3.wav	–0,715	–2,601
gcodcla.wav	1,781	–0,386
hcodryc.wav	2,291	–0,166
Hcodstr.wav	2,403	–0,128
Icodsna.wav	–3,029	–3,786
kcodsme.wav	3,093	0,038
lcodhrp.wav	1,041	–0,876
lcodpip.wav	1,973	–0,293
mcodcla.wav	–0,436	–2,331
ncodsfe.wav	3,135	0,045
scodclv.wav	1,689	–0,435

[1)] Наименования соответствующих параметров эталонных сигналов получаются путем замены подстроки «cod» в наименованиях параметров испытуемых сигналов на «ref», например наименованием параметра эталонного сигнала для «bcodtri.wav» является «breftri.wav».

Т а б л и ц а 6 — Параметры испытуемого сигнала и полученные значения DI и ODG для расширенной версии

Параметр	DI	ODG
Acodsna.wav	1,632	–0,467
Bcodtri.wav	2,000	–0,281
Ccodsax.wav	0,567	–1,300
Dcodryc.wav	1,725	–0,415
Ecodsmg.wav	1,594	–0,489
Fcodsb1.wav	1,039	–0,877
Fcodtr1.wav	1,555	–0,512
Fcodtr2.wav	0,162	–1,711
Fcodtr3.wav	–0,783	–2,662
Gcodcla.wav	1,457	–0,573
Hcodryc.wav	2,410	–0,126
Hcodstr.wav	2,232	–0,187
Icodsna.wav	–2,510	–3,664
Kcodsme.wav	2,765	–0,029
Lcodhrp.wav	1,538	–0,523
Lcodpip.wav	2,149	–0,219
Mcodcla.wav	0,430	–1,435
Ncodsfe.wav	3,163	0,050
Scodclv.wav	1,972	–0,293

6.2 Частота дискретизации и разрешение квантования

6.2.1 Параметр, подлежащий оценке

Оценивают частоту дискретизации и пропускную способность эталонного и обработанного аудиосигнала.

6.2.2 Метод оценки

Частота дискретизации зависит от полосы пропускания аудиосигналов. Для высококачественных аудиосигналов используется частота дискретизации 48 кГц. Должны быть выбраны значения частоты дискретизации и ширины полосы эталонного и обработанного аудиосигналов.

Разрешение квантования относится к динамическому диапазону аудиосигналов или шума квантования. Для высококачественных аудиосигналов используют линейный (или равномерный) метод квантования, имеющий 16-битное разрешение квантования. Значение разрешения и метод квантования должны быть идентифицированы.

6.2.3 Представление результатов оценки

Должны быть зарегистрированы выбранные и идентифицированные значения.

6.3 Задержка

6.3.1 Параметр, подлежащий оценке

Оценивают время задержки аудиосигналов от аудиовходов до кодера и их принятия, измеренное в секундах.

6.3.2 Метод оценки

Импульсные аудиосигналы должны быть использованы в качестве входных для позиции 2 рисунка 2. Время обработки сигнала между входом позиции 3 и выходом позиции 4 рисунка 2 должно быть измерено в секундах.

Примечание — Большинство систем аудиосвязи посредством цифровых сетей включает в себя схему буферизации. Следовательно, время буферизации также учитывают при измерении.

6.3.3 Представление результата оценки

Регистрируют измеренное время задержки в секундах.

7 Качество видео- и аудиосигналов в целом

7.1 Синхронизация аудио- и видеосигнала (синхронизация изображения и речевых сигналов)

7.1.1 Параметр, подлежащий оценке

Оценивают временну́ю синхронизацию между аудио- и видеоканалами.

7.1.2 Метод оценки

Полноценные мультимедийные системы в отличие от простого набора несвязанных медиаканалов способны поддерживать временну́ю синхронизацию между различными каналами. Следовательно, включение измерения качества временно́й синхронизации в показатели оценки качества систем аудио-, видеосвязи имеет чрезвычайно большое значение.

Алгоритм измерения временно́й синхронизации между медиаканалами приведен в Рекомендации ITU-T P.931. Он основан на исходном условии, что медиасигнал визуального канала может быть зарегистрирован на таких интерфейсах, как выход камеры и вход дисплея для видеоканала, и выход микрофона и вход громкоговорителя — для аудиоканала. Это исходное условие приведено на рисунке 1.

Медиасигналы, зарегистрированные такими интерфейсами, оцифровывают, при необходимости разбивают на кадры фиксированного размера и дают временные метки. Для получения дополнительной информации об этой процедуре см. Рекомендацию ITU-T P.931.

Оцифрованным кадрам аудио- и видеомедиапотоков присваивают порядковые номера следующим образом:

- A(m) и V(n) — входные аудио- и видеокадры, соответственно (m) и (n) — порядковые номера для каждого потока. Предполагается, что они связаны между собой и соответствуют одному и тому же событию;
- A′(p) и V′(q) — выходные аудио- и видеокадры соответственно;
- $T_A(m)$ и $T'_A(n)$ — временные метки для A(m) и A(n) соответственно. Временные метки для других кадров определяют аналогичным образом.

Для каждого входного кадра необходимо найти соответствующий выходной кадр, а не использовать все входные кадры, как описано в ITU-T P.931. Поскольку данные медиапотока изменяются, искажаются, пропускаются и переформируются, процесс согласования не является простым. Для видеокадров используют метод, применяющий параметры PSNR, рассмотренные в разделе 5. Для аудиокадров используют двухэтапный процесс, применяющий аудиоогибающие для этапа грубого согласования и спектральные плотности мощности — для этапа точного согласования. Дополнительная информация приведена в ITU-T P.931.

При таком рассмотрении предполагается, что между A(m) и A′(p), A′(p) и V′(n) и V′(q) установлены соотношения согласования. При этом предположении временно́й перекос между аудио- и видеокадрами рассчитывают по следующей формуле:

$$S_a\,(p,q) = Q'_{AV}(p,q) - Q_{AV}(m,n), \tag{8}$$

где $Q_{AV}(m,n) = T_A(m) - T_V(n)$ и $Q'_{AV}(p,q) = T'_A(p) - T'_V(q)$.

Примечание 1 — Для получения достоверного и значимого результата оценки важно выбрать подходящие входные аудиосигналы. Если видеосигнал содержит статические или близкие к статическим сцены, процесс согласования входных и выходных кадров будет затруднен или даже невозможен. Аналогичную осторожность следует соблюдать и при оценке аудиоканала.

Примечание 2 — Современные схемы сжатия видеосигнала дают колебания времени сжатия, передачи (при использовании кодирования с переменной скоростью передачи данных) и распаковки в зависимости от свойств входных сигналов. Следовательно, для оценки следует использовать подходящие входные сигналы, пригодные для предполагаемого применения.

Примечание 3 — Для систем с низкой частотой видеокадров иногда предпочтительнее иметь больший временно́й перекос между видео- и аудиопотоками, поскольку время задержки видеосигнала изменяется, в то время как аудиоданные обычно передаются изохронно.

Выбор стандартных входных аудио- и видеопотоков, пригодных для общего использования, оставлен для дальнейшего изучения.

7.1.3 Представление результатов оценки

Протокол измерений должен быть представлен в таком виде, чтобы любое различие между отдельными измерениями было четко проиллюстрировано. Также может быть представлена типичная сводная статистика (например, минимальное, максимальное, среднее и стандартное отклонение).

7.2 Масштабируемость

7.2.1 Параметр, подлежащий оценке

Оценивают автономную функцию динамической настройки частоты кадров в зависимости от доступной полосы пропускания между передатчиком и приемником.

7.2.2 Метод оценки

Метод измерения масштабируемости находится на рассмотрении.

7.2.3 Представление результатов оценки

Находится на рассмотрении.

7.3 Общее качество испытуемых сигналов

7.3.1 Параметр, подлежащий оценке

Оценивают коэффициент общего качества, являющийся характеристикой взаимодействия аудио- и видеосигнала.

7.3.2 Метод оценки

Общее качество систем аудио-, видеосвязи OQ_{AV} рассчитывают по формуле

$$OQ_{AV} = aQ_V + bQ_A + cQ_{V\&A}, \quad (9)$$

где Q_V — объективный показатель качества, оцененный в разделе 5;

Q_A — объективный показатель качества, оцененный в разделе 6;

$Q_{V\&A}$ — объективный показатель качества, оцененный в настоящем разделе.

a, b и c — коэффициенты взвешивания, зависящие от фактических применений системы аудио-, видеосвязи.

7.3.3 Представление результатов оценки

Должен быть представлен общий коэффициент качества с достаточной информацией об оцениваемой системе аудио-, видеосвязи.

Приложение А
(справочное)

PSNR, определенные в трехмерных пространствах, применяемые к гипотетическому ухудшению испытуемых сигналов по отношению к исходным эталонным видеосигналам

А.1 Введение

Настоящее приложение предназначено для демонстрации определений PSNR в трехмерном векторном пространстве для каждого пикселя, входящего в состав кадров видеосигналов. Определение PSNR в цветовом пространстве CIELAB приведено в уравнении (5), PSNR в пространстве sYCC — в уравнении (6), а PSNR в пространстве sRGB — в уравнении (4). Средняя цветовая разность, определенная в уравнении (1), также включена в это приложение для сравнения совместно с одномерными PSNR в пространствах *L** и *Y*.

Значения объективных показателей качества легко сравниваются с другими возможными показателями качества, которые могут быть применены в будущем, и результатами субъективной оценки качества видеосигнала.

А.2 Испытуемые сигналы и гипотетическое ухудшение

В настоящем приложении приведены 16 известных различных гипотетических ухудшений по сравнению с цифровыми видеофайлами, подготовленными в формате, установленном ITU-R BT.601-5, и используемыми при проведении оценки группой экспертов по качеству видеосигнала (VQEG). Исходные эталонные видеосигналы имеют обозначения от SRC13_REF__525.yuv до SRC22_REF__525.yuv, как показано в таблице А.1. Они используются с разрешения VQEG.

Программное обеспечение для различных объективных показателей качества видеосигналов было разработано в университете Чиба, Япония, в сотрудничестве с компанией Mitsubishi Electric Corp. Значения были получены для кадра уменьшенного размера 320 x 240 пикселей на кадр, более чем для 260 кадров. В уравнении (4) применяют следующие значения показателей: *P*1 = 1, *P*2 = 260, *M*1 = 1, *M*2 = 240 и *N*1 = 1, *N*2 = 320. Результаты расчета приведены в таблицах А.2 — А.6.

Таблица А.1 — Исходные эталонные видеосигналы, доступные для объективной оценки качества

Обозначение	Наименование	Содержание
SRC13_REF__525	Balloon-pops	Пленка, насыщенный цвет, движение
SRC14_REF__525	New York 2	Маскирующий эффект, движение
SRC15_REF__525	Mobile & Calendar	Цвет, движение
SRC16_REF__525	Betes_pas_betes	Цвет, синтезируемый, движение, вырезка сцен
SRC17_REF__525	Le_point	Цвет, прозрачность, движение во всех направлениях
SRC18_REF__525	Autumn leaves	Цвет, пейзаж, масштабирование, движение падающей воды
SRC19_REF__525	Football	Цвет, движение
SRC20_REF__525	Sailboat	Близкое к статическому изображение
SRC21_REF__525	Susie	Цвет кожи
SRC22_REF__525	Tempete	Цвет, движение

Таблица А.2 — PSNR в различных цветовых пространствах и цветовая разность для видеосигналов SRC13 и SRC14

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
hrc1/src13	20,5	23,2	23,6	26,3	26,3	8,3	hrc6/src13	22,2	22,6	22,1	25,9	24,9	6,2
hrc2/src13	23,6	23,5	23,2	25,9	25,0	5,4	hrc7/src13	22,2	21,1	20,7	23,0	22,1	5,9
hrc3/src13	22,2	22,7	22,3	25,6	24,6	5,8	hrc8/src13	21,9	22,3	21,9	25,3	24,5	6,7
hrc4/src13	21,4	22,1	21,7	25,6	24,7	7,4	hrc9/src13	21,6	20,6	20,3	22,8	21,8	6,9
hrc5/src13	20,4	19,3	19,0	21,2	20,3	8,0	hrc10/src13	22,1	20,9	20,6	23,0	22,0	6,3

Окончание таблицы А.2

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
hrc11/src13	21,7	22,8	22,5	24,5	25,3	6,9	hrc6/src14	24,5	23,6	23,8	25,3	24,1	5,0
hrc12/src13	22,4	23,6	23,3	24,8	26,0	5,9	hrc7/src14	24,5	21,5	21,7	22,5	21,4	4,1
hrc13/src13	21,3	20,7	20,6	23,4	22,2	6,8	hrc8/src14	24,3	23,5	23,7	24,9	24,0	5,3
hrc14/src13	21,2	20,3	20,0	22,7	21,6	7,9	hrc9/src14	24,3	21,4	21,7	22,4	21,4	4,5
hrc15/src13	21,9	22,1	21,7	25,3	24,4	7,6	Hrc10/src14	24,3	21,4	21,6	22,5	21,4	4,4
hrc16/src13	22,1	22,8	22,3	25,8	25,2	7,0	Hrc11/src14	25,5	26,0	26,1	24,6	26,3	4,1
hrc1/src14	22,4	25,8	25,9	26,6	28,1	7,5	Hrc12/src14	26,0	26,2	26,4	24,8	26,4	3,7
hrc2/src14	25,7	24,3	24,5	25,4	24,3	4,9	Hrc13/src14	21,5	20,8	21,7	23,2	21,7	5,4
hrc3/src14	24,9	23,8	24,0	25,1	24,1	4,7	Hrc14/src14	23,9	21,3	21,6	22,4	21,3	5,3
hrc4/src14	24,6	23,9	24,0	25,4	24,3	5,5	Hrc15/src14	25,8	25,8	26,0	27,2	26,3	5,6
hrc5/src14	22,5	19,7	20,0	20,7	19,6	5,9	Hrc16/src14	26,0	26,0	26,2	27,4	26,5	5,3

Примечание 1 — Видеосигналы hrc16/src14 и т. д. соответствуют гипотетически ухудшенному видеосигналу (hrc16) по сравнению с исходным эталонным видеосигналом (src14) соответственно.

Примечание 2 — Все видеосигналы представлены в размере 320×240 пикселей с 24-битной глубиной цвета.

Таблица А.3 — PSNR в различных цветовых пространствах и цветовая разность для видеосигналов SRC15 и SRC16

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
hrc1/src15	11,8	13,6	13,1	20,7	19,5	24,8	hrc1/src16	20,3	21,6	21,8	23,8	25,7	9,5
hrc2/src15	17,0	18,5	18,4	24,2	23,1	10,8	hrc2/src16	27,1	28,1	28,0	31,1	32,0	4,4
hrc3/src15	15,1	16,7	16,5	23,1	21,7	13,2	hrc3/src16	29,2	29,0	28,9	31,0	31,9	2,4
hrc4/src15	13,6	15,0	14,5	23,0	21,2	18,7	hrc4/src16	22,9	23,7	23,6	28,3	28,3	6,0
hrc5/src15	14,6	15,4	15,2	19,8	18,9	15,9	hrc5/src16	21,7	22,0	21,9	24,8	25,5	6,0
hrc6/src15	14,0	15,4	15,0	23,0	21,3	17,3	hrc6/src16	23,5	24,2	24,0	28,7	28,7	5,1
hrc7/src15	16,1	17,0	16,9	21,1	20,3	12,1	hrc7/src16	22,8	22,9	22,8	25,7	26,4	4,4
hrc8/src15	14,0	15,3	15,0	22,6	20,9	17,6	hrc8/src16	23,4	24,2	24,0	28,4	28,5	5,2
hrc9/src15	15,9	16,6	16,5	20,7	19,7	13,1	hrc9/src16	22,8	22,7	22,7	25,5	26,1	4,6
hrc10/src15	16,4	17,3	17,2	22,1	21,0	11,9	hrc10/src16	24,9	25,8	25,4	28,9	30,3	3,8
hrc11/src15	15,8	17,2	17,0	22,9	22,0	12,8	hrc11/src16	25,4	27,5	27,3	27,8	31,6	3,8
hrc12/src15	16,0	17,5	17,3	23,3	22,7	12,0	hrc12/src16	25,7	27,9	27,6	28,0	32,2	3,5
hrc13/src15	15,4	16,2	16,1	21,5	19,7	14,6	hrc13/src16	23,3	23,5	23,6	29,1	29,5	4,3
hrc14/src15	15,6	16,1	16,0	20,6	19,2	14,3	hrc14/src16	22,9	22,6	22,6	25,2	25,4	5,2
hrc15/src15	15,7	16,1	16,0	21,1	19,3	15,4	hrc15/src16	23,7	23,3	23,5	26,0	26,2	5,8
hrc16/src15	15,7	16,3	16,2	21,6	19,9	14,9	hrc16/src16	23,9	23,5	23,7	26,2	26,5	5,6

Примечание 1 — Видеосигналы hrc16/src16 и т. д. соответствуют гипотетически ухудшенному видеосигналу (hrc16) по сравнению с исходным эталонным видеосигналом (src16) соответственно.

Примечание 2 — Все видеосигналы представлены в размере 320×240 пикселей с 24-битной глубиной цвета.

Таблица А.4 — PSNR в различных цветовых пространствах и цветовая разность для видеосигналов SRC17 и SRC18

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
hrc1/src17	15,8	19,2	19,2	20,8	23,3	16,7	hrc1/src18	18,3	21,0	20,7	23,2	25,6	10,2
hrc2/src17	20,2	23,2	23,6	26,6	26,9	9,2	hrc2/src18	22,8	24,8	24,5	28,0	28,7	6,0
hrc3/src17	20,2	23,2	23,3	26,2	27,1	8,3	hrc3/src18	22,4	24,2	23,8	27,7	28,0	6,5
hrc4/src17	18,6	21,2	21,6	25,2	25,0	11,1	hrc4/src18	18,1	20,4	19,7	26,6	26,9	9,9
hrc5/src17	18,0	20,1	20,5	22,7	23,0	11,8	hrc5/src18	18,9	20,1	20,0	21,7	22,6	9,0
hrc6/src17	18,5	20,8	21,1	24,9	24,7	10,0	hrc6/src18	19,3	21,6	21,0	27,2	27,4	8,4
hrc7/src17	19,8	21,8	22,1	24,5	24,9	8,6	hrc7/src18	20,3	21,5	21,5	22,8	23,7	7,2
hrc8/src17	18,1	20,5	20,8	24,3	24,2	10,9	hrc8/src18	19,5	21,7	21,2	26,8	27,2	8,4
hrc9/src17	18,6	20,6	20,9	23,4	23,7	10,3	hrc9/src18	20,4	21,5	21,6	22,8	23,6	7,4
hrc10/src17	19,7	21,9	22,2	24,8	25,2	8,9	hrc10/src18	21,6	23,0	22,8	25,3	26,0	6,5
hrc11/src17	18,1	20,4	20,7	23,2	23,9	10,8	hrc11/src18	21,5	24,4	24,0	26,7	29,8	6,4
hrc12/src17	19,0	21,4	21,8	24,1	25,1	9,3	hrc12/src18	21,7	24,5	24,1	26,7	30,1	6,1
hrc13/src17	16,9	18,6	19,0	22,0	22,0	13,2	hrc13/src18	21,9	24,0	23,6	27,7	27,9	6,9
hrc14/src17	18,2	20,3	20,6	23,4	23,4	11,6	hrc14/src18	21,3	22,8	22,6	25,2	25,7	7,2
hrc15/src17	17,8	20,0	20,4	23,0	23,0	13,4	hrc15/src18	21,6	23,7	23,3	29,0	28,5	7,8
hrc16/src17	18,2	20,7	21,2	23,8	24,0	12,5	hrc16/src18	21,7	23,9	23,4	29,4	29,2	7,4

Примечание 1 — Видеосигналы hrc16/src18 и т. д. соответствуют гипотетически ухудшенному видеосигналу (hrc16) по сравнению с исходным эталонным видеосигналом (src18) соответственно.

Примечание 2 — Все видеосигналы представлены в размере 320×240 пикселей с 24-битной глубиной.

Таблица А.5 — PSNR в различных цветовых пространствах и цветовая разность для видеосигналов SRC19 и SRC20

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
hrc1/src19	20,0	22,6	22,6	23,2	25,6	7,8	hrc1/src20	15,8	17,4	17,2	20,1	20,2	12,7
hrc2/src19	23,6	25,1	24,9	27,9	28,6	4,8	hrc2/src20	20,6	20,8	20,7	23,7	22,1	6,9
hrc3/src19	23,1	24,6	24,4	27,7	28,0	5,8	hrc3/src20	18,7	19,3	19,3	22,6	21,3	8,2
hrc4/src19	20,3	22,5	22,1	26,6	27,1	6,9	hrc4/src20	18,7	19,2	19,0	22,6	21,0	8,8
hrc5/src19	19,8	20,8	20,7	22,5	23,3	7,4	hrc5/src20	18,7	16,2	16,0	18,5	16,6	8,3
hrc6/src19	20,6	22,7	22,3	27,0	27,2	6,6	hrc6/src20	18,8	19,4	19,2	23,1	21,4	8,1
hrc7/src19	21,0	21,7	21,7	22,8	23,5	5,9	hrc7/src20	19,4	17,5	17,3	19,6	18,1	7,2
hrc8/src19	20,7	22,6	22,3	26,6	26,9	6,8	hrc8/src20	18,6	19,2	19,0	22,8	21,2	8,4
hrc9/src19	21,2	23,1	22,7	27,2	27,2	6,2	hrc9/src20	20,0	20,3	20,1	23,5	22,0	6,4
hrc10/src19	20,0	21,3	21,1	24,4	24,7	7,8	Hrc10/src20	20,3	18,8	18,6	21,3	19,5	6,5
hrc11/src19	21,4	23,6	23,3	25,6	27,7	6,3	Hrc11/src20	19,9	21,5	21,4	23,3	23,8	6,6
hrc12/src19	22,4	24,7	24,4	25,9	28,8	5,4	Hrc12/src20	20,3	21,9	21,7	23,4	24,1	6,2
hrc13/src19	20,8	21,8	21,7	24,2	24,3	6,9	Hrc13/src20	19,8	18,8	18,7	21,4	19,8	7,8
hrc14/src19	21,1	22,1	22,0	24,8	25,1	7,1	Hrc14/src20	19,6	18,4	18,2	21,1	19,3	7,6
hrc15/src19	23,3	24,6	24,4	28,8	28,5	5,9	Hrc15/src20	19,5	20,3	20,2	23,5	22,2	8,4
hrc16/src19	23,6	25,1	24,8	29,4	29,4	5,4	Hrc16/src20	19,6	20,4	20,4	23,7	22,4	8,3

Примечание 1 — Видеосигналы Hrc16/src20 и т. д. соответствуют гипотетически ухудшенному видеосигналу (hrc16) по сравнению с исходным эталонным видеосигналом (src20) соответственно.

Примечание 2 — Все видеосигналы представлены в размере 320×240 пикселей с 24-битной глубиной.

Таблица А.6 — PSNR в различных цветовых пространствах и цветовая разность для видеосигналов SRC21 и SRC22

Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE	Обозначение	Lab	sYCC	sRGB	L*	Y	ΔE
Hrc1/src21	23,1	25,3	25,8	22,8	25,8	5,9	hrc1/src22	14,6	18,0	17,6	22,3	24,1	16,8
Hrc2/src21	29,3	29,1	29,2	28,5	29,6	3,2	hrc2/src22	18,9	21,7	21,0	26,3	26,5	9,3
Hrc3/src21	29,4	28,8	28,8	28,4	29,3	2,9	hrc3/src22	17,0	19,9	19,3	24,8	24,9	11,0
Hrc4/src21	28,4	27,7	27,9	27,3	28,2	3,5	hrc4/src22	17,4	20,1	19,4	25,4	25,6	11,2
Hrc5/src21	25,7	24,0	24,1	22,8	24,0	3,5	hrc5/src22	17,4	18,9	18,0	21,5	21,9	11,2
Hrc6/src21	29,5	28,3	28,5	27,9	28,6	2,8	hrc6/src22	17,2	20,0	19,3	25,7	25,8	10,8
Hrc7/src21	26,0	24,4	24,5	23,1	24,4	3,0	hrc7/src22	18,0	19,9	19,2	22,8	23,3	9,8
Hrc8/src21	29,1	28,1	28,3	27,5	28,4	3,0	hrc8/src22	17,2	19,9	19,2	25,1	25,2	11,1
Hrc9/src21	30,7	29,4	29,5	28,5	29,6	2,0	hrc9/src22	17,9	20,5	19,8	25,4	25,5	9,7
hrc10/src21	28,5	26,9	27,0	25,8	26,9	2,5	hrc10/src22	18,2	20,3	19,5	23,9	24,2	9,7
hrc11/src21	28,8	30,6	30,7	26,7	31,0	2,4	hrc11/src22	18,0	20,8	20,3	24,4	25,6	10,0
hrc12/src21	28,9	30,8	30,9	26,7	31,2	2,2	hrc12/src22	18,3	21,3	20,7	24,9	26,5	9,3
hrc13/src21	27,4	25,8	25,9	25,0	25,9	3,2	hrc13/src22	16,9	18,9	18,5	22,7	22,6	12,2
hrc14/src21	28,2	26,7	26,8	25,7	26,8	2,9	hrc14/src22	17,8	19,7	19,0	23,2	23,2	11,0
hrc15/src21	30,5	30,4	30,5	30,3	31,1	3,2	hrc15/src22	17,8	20,2	19,8	24,4	23,9	12,0
hrc16/src21	30,6	30,5	30,6	30,4	31,2	3,2	hrc16/src22	18,1	20,8	20,3	25,4	25,1	11,3

Примечание 1 — Видеосигналы hrc16/src22 и т. д. соответствуют гипотетически ухудшенному видеосигналу (hrc16) по сравнению с исходным эталонным видеосигналом (src22) соответственно.

Примечание 2 — Все видеосигналы представлены в размере 320×240 пикселей с 24-битной глубиной.

Приложение В
(справочное)

Сквозная объективная оценка качества видеосигнала в пространственно-частотной области

В.1 Параметр, подлежащий оценке

Оценивают среднеквадратические ошибки между соответствующими блоками в вейвлет-преобразованной области, соответствующей эталонному и ухудшенному видеосигналам, приведенные в ITU-R 6Q/42-E. Предполагается трехуровневое вейвлет-преобразование, поэтому существует 10 блоков, как показано на рисунках В.1 и В.2.

1
2
5
8

3
4

6
7

9
10

Рисунок В.1 — Назначение номеров блоков



Рисунок В.2 — Пример визуализированного вейвлет-разложения

В.2 Метод оценки

В качестве исходного эталонного сигнала для позиции 1 рисунка 2 используют эталонные видеосигналы, приведенные в таблице А.1. Видеосигналы с уменьшенным размером кадра в несжатом формате AVI должны быть подготовлены для позиции 2 рисунка 2. В этой точке необходимо встроить номера кадров, чтобы их можно было использовать для идентификации принятых кадров, соответствующих переданным кадрам.

Закодированные и переданные потоковые видеосигналы должны непрерывно приниматься. Должно быть проведено попиксельное вычисление.

Среднеквадратичные ошибки между каждым из соответствующих блоков $p = 1...10$ в исходном и ухудшенном видеокадрах k должны быть получены следующим образом:

Обозначим коэффициенты в вейвлет-области для позиции (i, j) блока p эталонных красных, зеленых и синих пиксельных данных как $C_{Ro,ijpk}$, $C_{Go,ijpk}$ и $C_{Bo,ijpk}$, соответственно и $C_{Rd,ijpk}$, $C_{Gd,ijpk}$ и $C_{Bdet,ijk}$ — для позиции (i, j) блока p ухудшенных красных, зеленых и синих пиксельных данных, соответственно.

Ухудшение d_{pk} в блоке p кадра k в вейвлет-области должно быть оценено как сумма квадратов ошибок, как приведено в формулах (В.1) и (В.2).

$$d_{pk} = \sum_i \sum_j \left(\Delta C^2_{Rijpk} + \Delta C^2_{Gijpk} + \Delta C^2_{Bijpk} \right), \quad \text{(В.1)}$$

где $\Delta C_{Rijpk} = C_{R\,d,ijpk} - C_{R\,o,ijpk}$

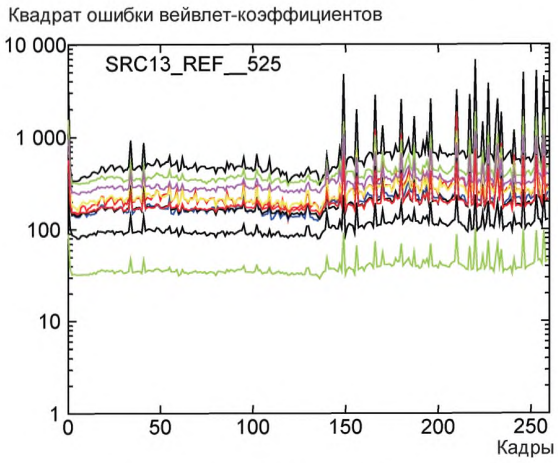
$$\Delta C_{Gijpk} = C_{G\,d,ijpk} - C_{G\,o,ijpk} \quad \text{(В.2)}$$

$$\Delta C_{Bijpk} = C_B - C_{B\,o,ijpk}$$

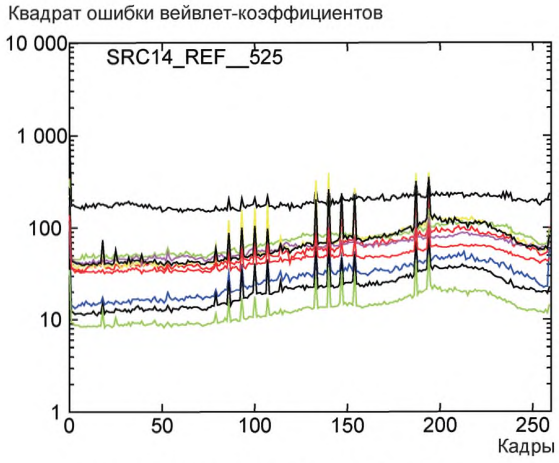
В.3 Представление результатов оценки

Должен быть построен график зависимости параметра суммы квадратов ошибок между блоками, соответствующими вейвлет-преобразованным кадрам, от номеров кадров, как показано на рисунке В.3, совместно с иден-

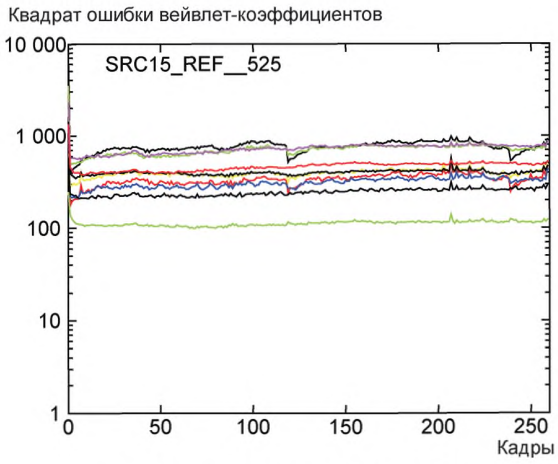
тификацией исходных эталонных видеосигналов. Также должны быть представлены условия измерения, такие как размер кадра в пикселях, частота кадров, потоковая скорость передачи данных.



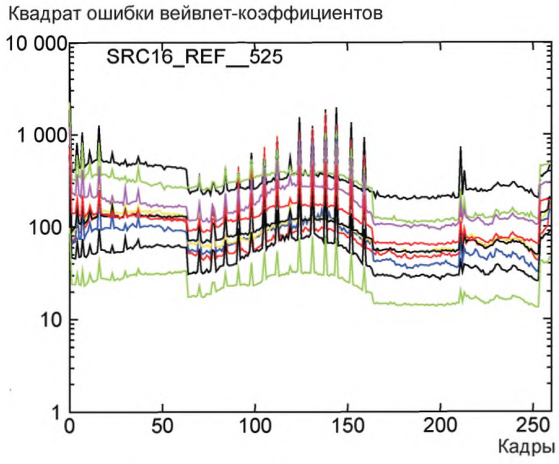
a) Пример для SRC13_REF__525



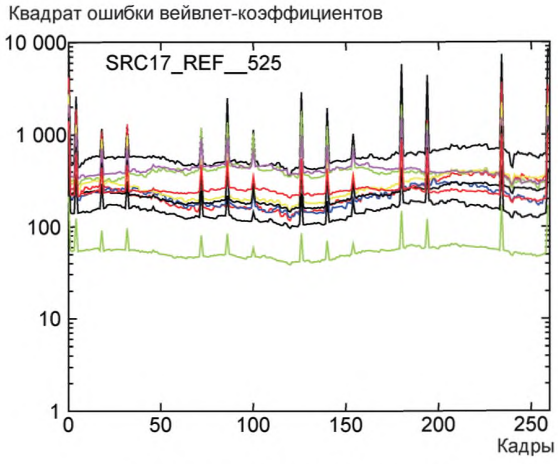
b) Пример для SRC14_REF__525



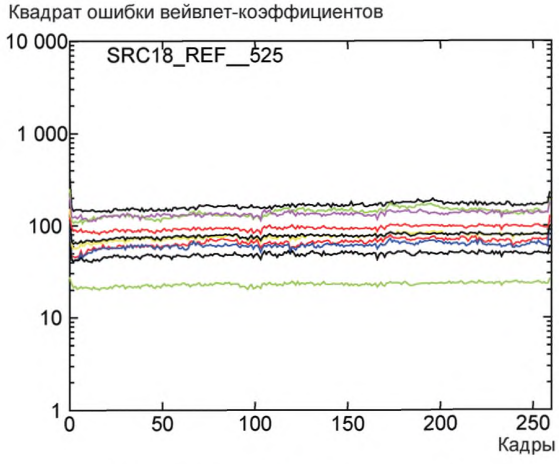
c) Пример для SRC15_REF__525



d) Пример для SRC16_REF__525

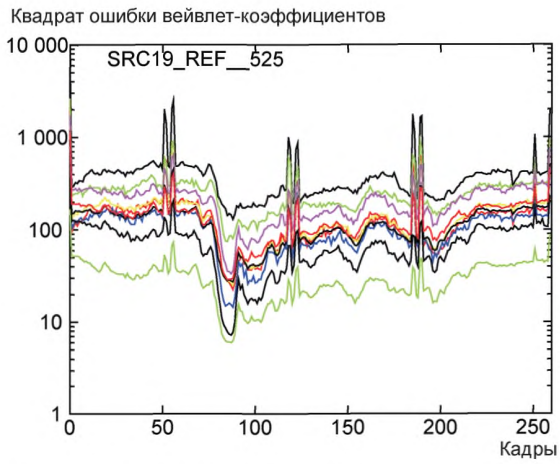


e) Пример для SRC17_REF__525

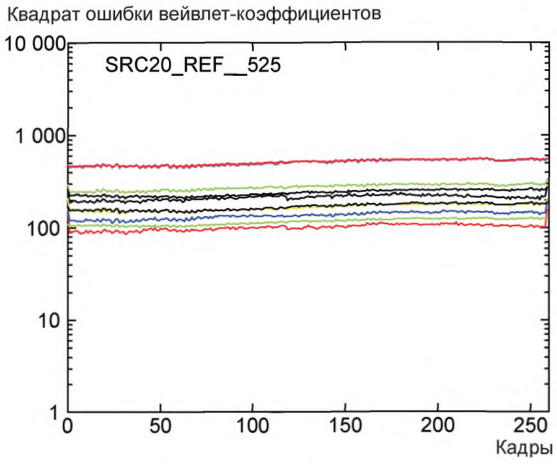


f) Пример для SRC18_REF__525

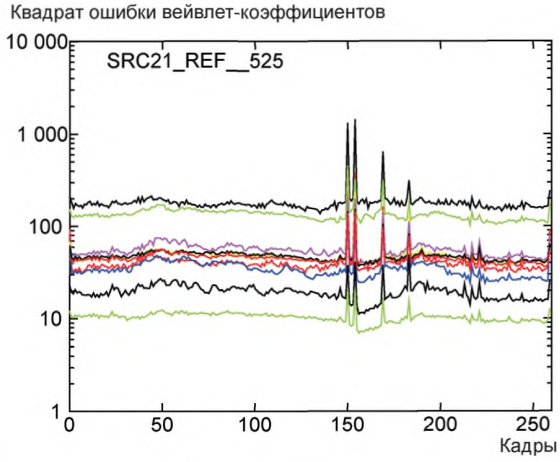
Рисунок В.3 — Диаграммы разности коэффициентов вейвлет-преобразования между эталонным и потоковым видеокадрами при скорости 250 кбит/с и 30 кадр/с, лист 1



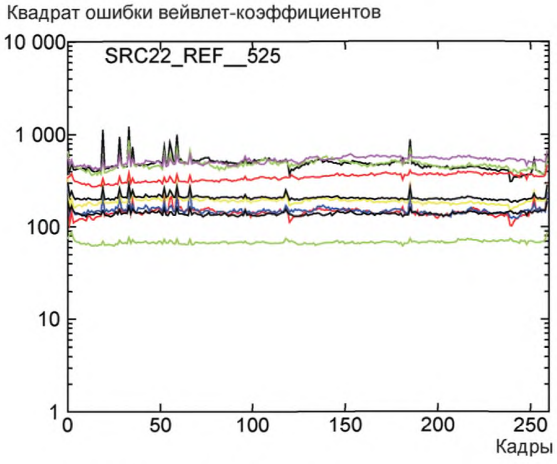
g) Пример для SRC19_REF__525



h) Пример для SRC20_REF__525



i) Пример для SRC21_REF__525



j) Пример для SRC22_REF__525

Условия оценки:
- размер видеокадра: 320 пикселей x 240 пикселей;
- частота кадров: 30 кадр/с;
- потоковая скорость передачи данных: 250 кбит/с;
- пропускная способность сети: более 250 кбит/с;
- воспроизведение: плейер Microsoft Media Player ® версия 7.1

Рисунок В.3, лист 2

Для получения результата оценки полученные квадраты ошибок также должны быть усреднены по кадрам согласно формуле (В.3), чтобы обеспечить общие параметры для объективной оценки, которые должны быть представлены, как показано в таблице В.1.

$$\overline{C_p} = \frac{1}{(K_2 - K_1 + 1)} \sum_{k=K_1}^{K_2} d_{kp} \qquad (B.3)$$

В целях оценки рейтинга качества видеосигнала (VQR) в виде единого параметра для каждого из полученных видеосигналов взвешенная сумма параметров VQR, приведенная в таблице 5 и рассчитанная в соответствии с формулой (В.4), должна быть представлена в крайнем правом столбце таблицы В.1.

$$VQR = w_0 + \sum_{p=1}^{10} w_p \overline{C_p}, \qquad (B.4)$$

где w_0 — смещение, а w_p при p =1...10 — весовые коэффициенты для *VQR*, которые должны наилучшим образом коррелировать с DMOS для набора эталонных видеосигналов, указанных в ITU-R 10-11Q и ITU-R WP 6Q (см. ITU-R 10-11Q/54-E).

Таблица В.1 — Сводная таблица разности коэффициентов вейвлет-коэффициентов

Исходный эталонный видеосигнал	$\overline{C}_1$	$\overline{C}_2$	$\overline{C}_3$	$\overline{C}_4$	$\overline{C}_5$	$\overline{C}_6$	$\overline{C}_7$	$\overline{C}_8$	$\overline{C}_9$	$\overline{C}_{10}$	VQR
SRC13_REF__525	725	300	440	212	275	343	109	201	203	40	20,4
SRC14_REF__525	197	64	76	30	77	62	23	74	47	14	14,7
SRC15_REF__525	785	346	714	314	401	728	245	404	464	112	43,3
SRC16_REF__525	388	120	289	94	117	191	53	105	125	25	17,1
SRC17_REF__525	733	309	438	241	317	443	153	247	262	56	28,2
SRC18_REF__525	165	67	140	61	77	134	49	78	95	23	18,7
SRC19_REF__525	441	150	266	113	152	217	74	128	140	30	19,9
SRC20_REF__525	212	101	273	136	165	500	168	237	510	116	35,1
SRC21_REF__525	187	42	136	35	49	56	20	48	45	10	14,3
SRC22_REF__525	483	147	472	150	191	522	139	207	342	68	29,7

Примечание — Значения VQR напрямую зависят от набора применяемых весовых коэффициентов. Пример в крайнем правом столбце условно основан на наборе весовых коэффициентов, подготовленном в университете Чиба в январе 2002 года.

Приложение С
(справочное)

Описание метода объективного измерения PEAQ

С.1 Базовая структура алгоритма измерения PEAQ

Базовая структурная схема метода объективного измерения PEAQ приведена на рисунке С.1. Она состоит из двух входов, один из которых предназначен для (необработанного) эталонного сигнала, а другой — для испытуемого сигнала. Последний может быть, например, выходным сигналом кодека, на который подается эталонный сигнал.

Этот метод измерения применим к большинству типов оборудования для обработки аудиосигналов, как цифровых, так и аналоговых. Однако ожидается, что многие приложения будут ориентированы на аудиокодеки.

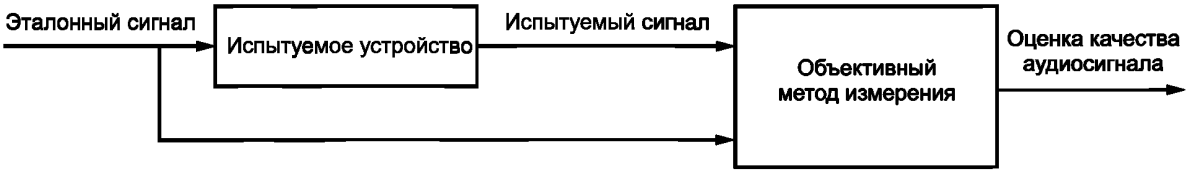


Рисунок С.1 — Базовая структурная схема проведения объективных измерений

Высокоуровневое представление модели PEAQ показано на рисунке С.2. Метод PEAQ основан на общепринятых психоакустических принципах. В целом он сравнивает сигнал, который был некоторым образом обработан, с соответствующим выровненным по времени эталонным сигналом. На первом этапе обработки сигнала моделируют периферийную акустическую модель («модель восприятия», или «модель уха»). Одновременные кадры эталонного и обработанного сигналов преобразуют в выходные сигналы акустических моделей. На следующем этапе алгоритма моделируют звуковое искажение, присутствующее в испытуемом сигнале, путем сравнения выходных сигналов акустических моделей. Информация, полученная посредством этих процессов, дает несколько значений, так называемых MOV (выходных переменных модели), и может использоваться для детального анализа сигнала.

Конечной целью является введение параметра качества, состоящего из одного числа, указывающего на слышимость искажений, присутствующих в испытуемом сигнале. Для сохранения этого параметра требуется некоторая дальнейшая обработка MOV, имитирующая когнитивную часть слуховой системы человека. Следовательно, алгоритм PEAQ использует искусственную нейронную сеть.

Существуют две версии PEAQ: базовая версия, отличающаяся низким уровнем сложности, и расширенная версия, обеспечивающая большую точность за счет более высокой сложности. Структура обеих версий очень похожа и точно соответствует модели PEAQ, показанной на рисунке С.2. Основное различие между базовой и расширенной версиями состоит в соответствующих моделях уха и набора, применяемых MOV. Базовая и расширенная версии описаны в разделах С.2 и С.3.

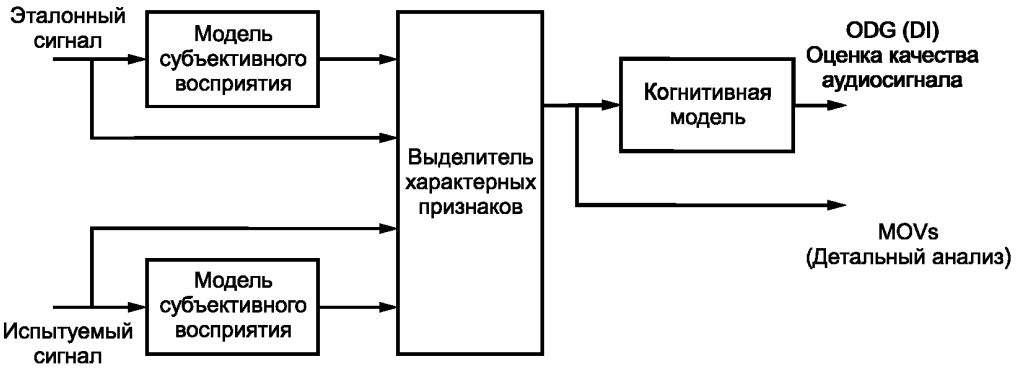


Рисунок С.2 — Представление модели PEAQ

С.2 Базовая версия

В базовой версии реализована акустическая модель на основе FFT, как показано на рисунке С.3.

Большинство характерных признаков этой модели основано на фундаментальных психоакустических принципах. На рисунке С.3 показано прохождение сигнала от входного сигнала до конечного расчета диаграммы возбуждения. Обработка начинается с преобразования входного сигнала в частотный домен (интервал).

Применяется 2048-точечное FFT с последующим масштабированием спектров в соответствии с уровнем прослушивания, значение которого должно быть введено пользователем в качестве параметра. Этот процесс дает частотное разрешение, составляющее приблизительно 23,4 Гц, и соответствующее временное разрешение 23,4 мс (при частоте дискретизации 48 кГц).

В конструктивном блоке эффекты акустической модели наружного и среднего уха моделируют путем взвешивания спектра с подходящими фильтрующими функциями. Затем спектры группируют в критические полосы, сохраняя разрешение 1/4 барка[1)] на полосу. Последующее добавление «внутреннего шума» предназначено для моделирования эффектов, таких как постоянная маскировка звуков в нашей слуховой системе, вызываемых циркуляцией крови и другими физиологическими явлениями. За этим этапом следует вычисление маскирующих эффектов. Одновременную маскировку моделируют функцией распределения, зависимой от частоты и уровня. Временная маскировка моделируется лишь частично, поскольку временное разрешение представляет собой тот же диапазон, что и синхронизация любых фоновых маскирующих эффектов, которые в результате этого не могут быть смоделированы. Эксперименты показали, что обратная маскировка очень грубо моделируется побочными эффектами FFT.

Используя устройство выделения характерных признаков, из компенсации выходного сигнала акустической модели извлекают одиннадцать MOV. В таблице С.1 приведены перечень этих MOV и их интерпретация. Дополнительная информация о MOV приведена в приложении ITU-R рекомендации BS 1387-1.

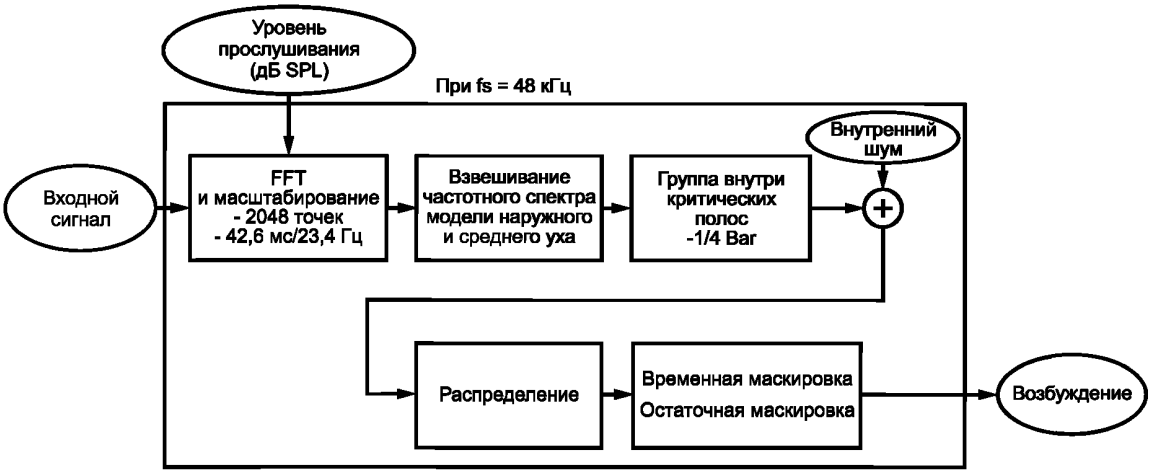


Рисунок С.3 — Акустическая модель на основе FFT, базовая версия PEAQ

Таблица С.1 — Переменные выходные данные модели, базовая версия PEAQ

Переменные выходные данные модели (MOV)	Назначение
$WinModDiff1_B$	Изменения в модуляции (связанные с резкостью)
$AvgModDiff1_B$	
$AvgModDiff2_B$	
$RmsNoiseLoud_B$	Громкость искажения
$BandwidthRef_B$	Линейные искажения (частотная характеристика и т. д.)
$BandwidthTest_B$	
$RelDistFrames_B$	Частота слышимых искажений
Total NMR_B	Отношение шума к маске
$MFPD_B$	Вероятность обнаружения
ADB_B	
EHS_B	Гармоническая структура ошибки

[1)] Барк (z) — психофизическая единица высоты звука.

С.3 Расширенная версия

В расширенной версии используют некоторые MOV, полученные путем реализации акустической модели базовой версии, однако в дополнение к этому она реализует вторую акустическую модель с улучшенным временным разрешением, как показано на рисунке С.4.

По сравнению с базовой версией эта модель выполняет частотно-временное преобразование с использованием набора фильтров, группируя сигнал в 40 полос слышимых частот с временным разрешением около 0,66 мс. Этот процесс позволяет с высокой степенью точности моделировать эффекты обратной маскировки. После расчета обратной и одновременной маскировки сигнал является субдискретизированным (повторно отобранным) с коэффициентом 1:6 для повышения эффективности расчета. После добавления к субдискретизированному сигналу внутреннего шума и конечного моделирования эффектов остаточной маскировки выход этой модели становится повторно возбужденным.

По сравнению с базовой версией на основе FFT улучшается временное разрешение, что позволяет лучше моделировать временные эффекты за счет частотного разрешения и сложности вычислений.

Благодаря сочетанию параметров, полученных из обеих акустических моделей, количество MOV, используемых расширенной версией для получения окончательного параметра качества, может быть уменьшено до пяти. При этом немного улучшается точность алгоритма по сравнению с базовой версией. MOV, используемые расширенной версией, приведены в таблице С.2. Дополнительная информация о расширенной версии приведена в приложении к ITU-R BS.1387-1.

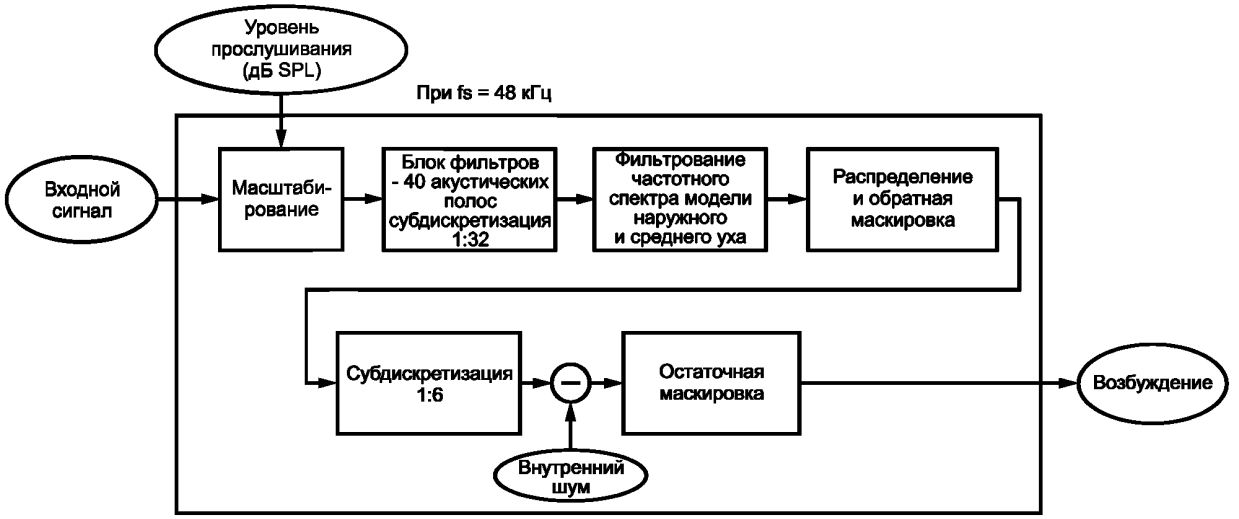


Рисунок С.4 — Акустическая модель на основе набора фильтров, расширенная версия PEAQ

Таблица С.2 — Переменные выходные данные модели, расширенная версия PEAQ

Выходные переменные данные модели (MOV)	Назначение
$RmsNoiseLoudAsym_A$	Громкость искажения
$RmsModDiff_A$	Изменения в модуляции (связанные с резкостью)
$AvgLinDist_A$	Линейные искажения (частотная характеристика и т. д.)
Segmental NMR_B	Отношение шума к маске
EHS_B	Гармоническая структура ошибки

С.4 Выходной параметр метода измерения PEAQ

Выходным параметром метода PEAQ является уровень объективной разности (ODG), соответствующий уровню субъективной разности (SDG) в субъективной области. Точность ODG ограничена одним десятичным знаком. Однако следует соблюдать предусмотрительность и в целом не ожидать, что разница между любой парой ODG в одну десятую значения уровня будет являться значительной. Это замечание справедливо и при рассмотрении результатов субъективного испытания на прослушивание. ODG также может принимать положительные значения.

Такие значения могут возникать, потому что PEAQ использует когнитивную модель для сопоставления MOV с результатами субъективного испытания на прослушивание. В случае субъективных испытаний на прослушивание SDG может принимать положительное значение в тех случаях, когда испытуемый неправильно установил эталонный и испытательный сигналы.

Показатель искажения (DI) имеет то же значение, что и ODG. Однако DI и ODG можно сравнивать только количественно, но не качественно. DI характеризуется меньшим насыщением, чем насыщение значения ODG. Кроме того, отличается диапазон значений. Как правило, ODG следует использовать в качестве показателя качества для значений ODG, превышающих приблизительно минус 3,6. ODG очень хорошо коррелирует с субъективной оценкой в этом диапазоне. Если значение ODG составляет меньше минус 3,6, следует использовать DI.

С.5 Эффективность метода измерения PEAQ

Эффективность модели PEAQ может зависеть от значений целого ряда различных показателей. Корреляция между ODG и SDG является очевидным показателем оценки эффективности. Кроме того, для проверки эффективности были использованы два дополнительных показателя, учитывающие надежность среднего значения — оценка абсолютных ошибок (AES) и схема устойчивости сигнала.

Проведенные ITU-R контрольные испытания показали, что PEAQ с высокой точностью предсказывает воспринимаемое качество и превосходит ранее существующие методы измерений. Дополнительная информация приведена в приложении к ITU-R BS.1387-1 и [AES-PEAQ][1)].

[1)] T. Theide и др. «ОВКА — Стандарт МСЭ для объективного измерения воспринимаемого качества аудиосигнала», J. Audio Eng. Soc., том 48, с. 3—29 (январь/февраль 2000 г.)

Приложение ДА
(справочное)

Сведения о соответствии ссылочных международных стандартов национальным стандартам

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
IEC 60268-4	—	*
IEC 60268-5	—	*
IEC 61146-1:1994	—	*
IEC 61146-2:1997	—	*
IEC 61966-2-1:1999	—	*
IEC 61966-3:2000	—	*
IEC 61966-4:2000	—	*
IEC 61966-5:2000	—	*
IEC 61966-9:2000	—	*
ITU-R BS.1387-1:2001	—	*
ITU-R BT.601-5:1995	—	*
ITU-T J.144:2001	—	*
ITU-T P.931:1998	—	*
* Соответствующий национальный стандарт отсутствует. До его принятия рекомендуется использовать перевод на русский язык данного международного стандарта. Официальный перевод данного международного стандарта находится в Федеральном информационном фонде стандартов.		

Библиография

ITU-R 10-11Q/56-E:2001 Canada (on behalf of the Entire VQEG body) — Draft Video Quality Experts Group's Results [Канада (от имени всех членов VQEG) — Предварительные результаты группы экспертов по качеству видеосигнала]

ITU-R 6Q/39-E:2001 Liaison Rapporteur with U.S. Committee T1A1, Documentation of objective video quality metrics (Докладчик по связям с Комитетом США T1A1, Документация по объективным параметрам качества сигнала для видеоматериалов)

ITU-R 6Q/42-E:2001 Republic of Korea — Proposed Preliminary Draft New Recommendation — A new method for objective measurement of video quality using wavelet transform (Республика Корея — Предлагаемый предварительный проект новой рекомендации. Новый метод объективного измерения качества видеосигнала с использованием вейвлет-преобразования)

ITU-T P.930:1996 Principles of a reference impairment system for video (Принципы системы ухудшения качества передачи эталонного видеосигнала)

ITU-T G.113:2001 Transmission impairments due to speech processing, Appendix I: Provisional planning values for the equipment impairment factor Ie and packet-loss robustness factor Bpl (Ухудшения качества передачи из-за обработки речи. Приложение I. Предварительные запланированные значения для коэффициента ухудшения качества оборудования Ie и коэффициента устойчивости к потере пакетов Bpl)

ITU-T P.862:2001 Objective quality measurement of telephone-band (300—3400 Hz) speech codecs (Объективное измерение качества речевых кодеков диапазона телефонной связи (300—3400 Гц)

T. Theide *et.al.* PEAQ — The ITU standard for Objective Measurement of Perceived Audio Quality, J. Audio Eng. Soc., vol.48, pp 3—29 (2000 Jan./Feb.) (PEAQ — Стандарт МСЭ для объективного измерения воспринимаемого качества аудиосигнала)

Measuring quality in videoconferencing systems, Part number PC316, Intel Corporation (November 1997) (Измерение качества в системах видеоконференцсвязи, каталожный номер PC316, Intel Corporation)

Criteria for product evaluation, NASA Desktop video expert center, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California (August 1997) (Критерии оценки продукта, экспертный центр по настольным видеосистемам NASA)

Quality aspects of computer-based video services, Norbert Gerfelder (Fraunhofer Institute for Computer Graphics, Darmstadt, Germany and Wolfgang Muller (Darmstadt Technical University), (Oct. 1995) (Качественные аспекты компьютерных видеоуслуг)

Comparative study on narrow-bandwidth presentation of streaming educational videos, H. Ikeda, S. Dickerson, Y. Higaki, Journal of Faculty of Engineering, Chiba University, Vol. 49,No. 1, pp.19—26 (1997-9) (Сравнительное исследование узкополосного представления потоковых образовательных видеоматериалов)

УДК 621.377:006.354 ОКС 33.160.60 ОКПД2 26.30

Ключевые слова: аудиосигнал, видеосигнал, эталонный видеосигнал, модель, метод оценки, оценка качества, сквозное воспроизведение цвета, сквозное воспроизведение тона, PEAQ, DMOS, PSNR, VQR

БЗ 5—2018/47

Редактор *Л.В. Коретникова*
Технический редактор *В.Н. Прусакова*
Корректор *Л.С. Лысенко*
Компьютерная верстка *Е.О. Асташина*

Сдано в набор 18.10.2018. Подписано в печать 14.11.2018. Формат 60×84$^{1}/_{8}$. Гарнитура Ариал.
Усл. печ. л. 4,65. Уч.-изд. л. 4,21.
Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении ФГУП «СТАНДАРТИНФОРМ» для комплектования Федерального информационного фонда стандартов, 117418 Москва, Нахимовский пр-т, д. 31, к. 2.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru