

МСЭ-R

Сектор радиосвязи МСЭ

Рекомендация МСЭ-R ВТ.1908-0
(01/2012)

**Методы объективного измерения
качества изображения для
радиовещательных применений с
использованием ТВЧ при наличии
эталонного сигнала с ухудшенными
характеристиками**

Серия ВТ
Радиовещательная служба
(телевизионная)



Предисловие

Роль Сектора радиосвязи заключается в обеспечении рационального, справедливого, эффективного и экономичного использования радиочастотного спектра всеми службами радиосвязи, включая спутниковые службы, и проведении в неограниченном частотном диапазоне исследований, на основании которых принимаются Рекомендации.

Всемирные и региональные конференции радиосвязи и ассамблеи радиосвязи при поддержке исследовательских комиссий выполняют регламентарную и политическую функции Сектора радиосвязи.

Политика в области прав интеллектуальной собственности (ПИС)

Политика МСЭ-R в области ПИС излагается в общей патентной политике МСЭ-Т/МСЭ-R/ИСО/МЭК, упоминаемой в Резолюции МСЭ-R 1. Формы, которые владельцам патентов следует использовать для представления патентных заявлений и деклараций о лицензировании, представлены по адресу: <http://www.itu.int/ITU-R/go/patents/en>, где также содержатся Руководящие принципы по выполнению общей патентной политики МСЭ-Т/МСЭ-R/ИСО/МЭК и база данных патентной информации МСЭ-R.

Серии Рекомендаций МСЭ-R

(Представлены также в онлайн-форме по адресу: <http://www.itu.int/publ/R-REC/en>.)

Серия	Название
BO	Спутниковое радиовещание
BR	Запись для производства, архивирования и воспроизведения; пленки для телевидения
BS	Радиовещательная служба (звуковая)
BT	Радиовещательная служба (телевизионная)
F	Фиксированная служба
M	Подвижные службы, служба радиоопределения, любительская служба и относящиеся к ним спутниковые службы
P	Распространение радиоволн
RA	Радиоастрономия
RS	Системы дистанционного зондирования
S	Фиксированная спутниковая служба
SA	Космические применения и метеорология
SF	Совместное использование частот и координация между системами фиксированной спутниковой службы и фиксированной службы
SM	Управление использованием спектра
SNG	Спутниковый сбор новостей
TF	Передача сигналов времени и эталонных частот
V	Словарь и связанные с ним вопросы

Примечание. – Настоящая Рекомендация МСЭ-R утверждена на английском языке в соответствии с процедурой, изложенной в Резолюции МСЭ-R 1.

Электронная публикация
Женева, 2020 г.

© ITU 2020

Все права сохранены. Ни одна из частей данной публикации не может быть воспроизведена с помощью каких бы то ни было средств без предварительного письменного разрешения МСЭ.

РЕКОМЕНДАЦИЯ МСЭ-R ВТ.1908-0*

Методы объективного измерения качества изображения для радиовещательных применений с использованием ТВЧ при наличии эталонного сигнала с ухудшенными характеристиками

(2012)

Сфера применения

В настоящей Рекомендации определяются методы оценки воспринимаемого качества видеоизображения для радиовещательных применений с использованием ТВЧ при наличии эталонного сигнала с ухудшенными характеристиками.

Ассамблея радиосвязи МСЭ,

учитывая,

- a)* что возможность автоматически измерять качество вещательного видеосигнала уже давно признана в качестве ценного ресурса для промышленности;
- b)* что в Рекомендации МСЭ-R ВТ.1683 описываются объективные методы измерения воспринимаемого качества видеоизображения для цифрового телевизионного радиовещания стандартной четкости при наличии ухудшенного эталонного сигнала;
- c)* что в Рекомендации МСЭ-R ВТ.709 описываются значения параметров стандартов ТВЧ для производства и международного обмена программами, а в Рекомендации МСЭ-R ВТ.500 описываются методы субъективной оценки качества видеоизображения, включая телевидение высокой четкости;
- d)* что ТВЧ все более широко используется в радиовещании;
- e)* что 9-я Исследовательская комиссия МСЭ-Т, по итогам доклада о ТВЧ, присланного VQEG, разработала Рекомендацию МСЭ-Т J.342, в которой описывается объективное измерение качества видеоизображения ТВЧ при наличии эталонного сигнала с ухудшенными характеристиками;
- f)* что объективное измерение воспринимаемого качества видеоизображения ТВЧ может дополнять методы субъективной оценки,

рекомендует,

1 чтобы для объективного измерения воспринимаемого качества видеоизображения в радиовещательных применениях с использованием ТВЧ при наличии эталонного сигнала с ухудшенными характеристиками использовалась модель объективного качества видеоизображений, приведенная в Приложении 1.

* В феврале 2020 года 6-я Исследовательская комиссия по радиосвязи внесла поправки редакционного характера в настоящую Рекомендацию в соответствии с Резолюцией МСЭ-R 1.

Приложение 1

1 Введение

В настоящей Рекомендации представлен метод измерения качества видеоизображений для использования в телевидении высокой четкости (ТВЧ) без интерактивных приложений в тех случаях, когда может быть применен метод измерения при наличии эталона с ухудшенными характеристиками (RR). Оценки по этой модели сравнивались с субъективными оценками качества, полученными с использованием Рекомендации МСЭ-R ВТ.500. Анализы показали, что точность этой модели аналогична точности при использовании пикового отношения сигнал/шум (PSNR).

Для того чтобы модель RR работала правильно, в ее распоряжении должен иметься неискаженный источник видеоизображения для извлечения необходимых параметров. Эти извлеченные параметры, а также искаженная видеопоследовательность являются входными данными для модели RR. Метод оценки выполняет как калибровку (т. е. регистрацию усиления/смещения и пространственных/временных характеристик), так и объективную оценку качества видеоизображения.

Материалы проверочных тестов содержали данные о снижении эффективности кодирования как по методу ITU-T H.264, так и по методу MPEG-2 и о различных ситуациях возникновения ошибок передачи (битовые ошибки, потерянные пакеты). Модель, описанная в настоящей Рекомендации, может быть использована для контроля качества развертываемых сетей в целях обеспечения их эксплуатационной готовности. Визуальные эффекты ухудшений могут включать как пространственные, так и временные ухудшения. Модель, рассматриваемая в данной Рекомендации, также может быть использована для лабораторного тестирования видеосистем. В случае использования для сравнения различных видеосистем, целесообразно применять количественный метод (например, тот, который приведен Рекомендации МСЭ-T J.149), чтобы определить точность рассматриваемой модели для данного конкретного контекста.

Настоящая Рекомендация считается подходящей для использования в радиовещательных службах со скоростями передачи в диапазоне от 1 Мбит/с до 30 Мбит/с. В проверочном испытании были рассмотрены следующие величины разрешения и частоты кадров:

- 1080/59,94/I
- 1080/25/P
- 1080/50/I
- 1080/29,97/P.

В проверочном испытании для каждого уровня разрешения предполагались следующие условия:

Проверяемые факторы
Видеоразрешение: 1 920 × 1 080 чересстрочное и прогрессивное
Частота видеокадров 29,97 и 25 кадров в секунду
Скорости передачи видеоданных: от 1 до 30 Мбит/с
Временная остановка кадров (паузы с пропусками) максимум 2 секунды
Ошибки передачи с потерей пакетов
Преобразование канала SRC из 1 080 в 720/P, сжатие, передача, декомпрессия, а затем обратное преобразование в 1 080
Технологии кодирования
H.264/AVC (MPEG-4 Part 10)
MPEG-2

Следует отметить, что разрешение 720/P рассматривалось в плане проверочного испытания как часть условия тестирования (HRC). Поскольку в настоящее время 720/P обычно масштабируется как часть отображения, было высказано мнение, что условия HRC 720/P будут более подходящими для этого формата.

1.1 Применения

Применения моделей оценки, описанных в настоящей Рекомендации, включают, помимо прочего, следующие возможные варианты:

- 1) контроль качества видеоизображений в приемнике при наличии боковых каналов;
- 2) контроль качества видеоизображений на измерительных узлах, расположенных между точкой передачи и точкой приема.

Модель, описываемая в настоящей Рекомендации, дает статистически аналогичные показатели для PSNR, но еще она может быть использована для оценки качества видеоизображения при наличии эталонного сигнала с ухудшенными характеристиками в точке измерения.

1.2 Ограничения

Модель оценки качества видеоизображения, описываемая в настоящей Рекомендации, не может быть использована для замены субъективного тестирования. Коэффициенты корреляции между двумя тщательно разработанными и выполненными субъективными тестами (т. е. в двух разных лабораториях), обычно находятся в диапазоне от 0,95 до 0,98. Эта Рекомендация не может быть использована для сравнения видеосистем (например, для сравнения двух кодеков, сравнения двух различных реализаций одного и того же алгоритма сжатия). Характеристики модели оценки качества видеоизображения, описываемой в настоящей Рекомендации, статистически не лучше, чем при использовании PSNR.

При наличии статических кадров условия испытаний обычно предусматривали остановку кадров продолжительностью менее 2 секунд. Модель, приведенная в данной Рекомендации, не проверялась на предмет измерения качества видеоизображений в условиях повторной буферизации (т. е. видеопоследовательность с постоянно возрастающей задержкой или с остановками без пропусков). Эта модель не тестировалась для других частот кадров, отличающихся от тех, которые используются в телевизионных системах (например, 29,97 кадров в секунду и 25 кадров в секунду в чересстрочном или прогрессивном режимах).

Следует отметить, что при новых технологиях кодирования и передачи, создающих артефакты, которые не были включены в эту оценку, объективная модель может привести к ошибочным результатам. В таких случаях необходима субъективная оценка.

Отмечается, что модель, рассматриваемая в данной Рекомендации, не оценивалась на контенте "говорящих голов", типичном для сценариев видеоконференций.

2 Ссылки

Следующие Рекомендации МСЭ и другие ссылки содержат положения, которые посредством ссылок в этом тексте составляют положения настоящей Рекомендации. На момент публикации указанные издания были действующими. Все Рекомендации и другие ссылки подлежат пересмотру; поэтому пользователям этой Рекомендации рекомендуется изучить возможности получения самых последних изданий Рекомендаций и других материалов, ссылки на которые перечислены ниже. Перечень действующих в настоящее время Рекомендаций МСЭ регулярно публикуется.

Ссылка на документ, приведенная в данной Рекомендации, не дает статуса Рекомендации как отдельного документа.

Recommendation ITU-T J.244 (2008), *Full reference and reduced reference calibration methods for video transmission systems with constant misalignment of spatial and temporal domains with constant gain and offset.*

3 Определения

3.1 Термины, определенные в других местах

В настоящей Рекомендации используются следующие термины, определенные в других местах:

3.1.1 субъективная оценка (по изображениям) (subjective assessment (picture)): Определение качества или ухудшения в похожих на программу изображениях, представленные группе экспертов в ходе сеансов просмотра.

3.1.2 объективное измерение воспринимаемого качества (по изображениям) (objective perceptual measurement (picture)): Измерение характеристик программной цепи с помощью похожих на программу изображений и методов объективных (инструментальных) измерений для получения показаний, приближающихся к оценке, которая могла бы быть получена от субъективного оценочного испытания.

3.1.3 инициатор (proponent): Организация или компания, которая предлагает модель оценки качества видеоизображений для проверочного тестирования и возможного включения в какую-либо Рекомендацию МСЭ.

3.2 Термины, определенные в настоящей Рекомендации

В данной Рекомендации определяются следующие термины:

3.2.1 частота кадров (frame rate): Количество уникальных кадров (т. е. общее количество кадров – количество повторяемых кадров) в секунду.

3.2.2 имитированные ошибки передачи (simulated transmission errors): Ошибки, наложенные на цифровой поток видеосигнала в строго контролируемой среде. Примеры включают моделированные коэффициенты потери пакетов и моделированные битовые ошибки. Параметры, используемые для контроля моделированных ошибок передачи четко определяются.

3.2.3 ошибки передачи (transmission errors): Любые ошибки, наложенные на видеоизображение при передаче. Примеры типов ошибок включают моделированные ошибки передачи и условия реальной сети.

4 Сокращения и акронимы

Эта Рекомендация использует следующие аббревиатуры и сокращения:

ACR	Absolute Category Rating (see Recommendation ITU-R BT.500)	Оценка по абсолютным категориям (см. Рекомендацию МСЭ-R ВТ.500)
ACR-HR	Absolute Category Rating with Hidden Reference (see Recommendation ITU-T P.910)	Оценка по абсолютным категориям со скрытым эталоном (см. Рекомендацию МСЭ-T P.910)
AVI	Audio Video Interleave	Чередование аудио и видео
DMOS	Difference Mean Opinion Score	Разностный показатель усредненного мнения
FR	Full Reference	Полный эталонный сигнал
FRTV	Full Reference TeleVision	Телевидение с полным эталонным сигналом
HRC	Hypothetical Reference Circuit	Гипотетическая эталонная цепь
ILG	VQEG's Independent Laboratory Group	Группа "Независимая лаборатория" VQEG
MOS	Mean Opinion Score	Показатель усредненного мнения
MOSp	Mean Opinion Score, predicted	Прогнозируемый показатель усредненного мнения
NR	No (or Zero) Reference	Нет эталона (или нулевой эталон)
PSNR	Peak Signal-to-Noise Ratio	Пиковое отношение сигнал/шум
PVS	Processed Video Sequence	Обработанная видеопоследовательность

RMSE	Root Mean Square Error	Среднеквадратическая ошибка
RR	Reduced Reference	Эталон с ухудшенными характеристиками
SFR	Source Frame Rate	Частота кадров источника
SRC	Source Reference Channel or Circuit	Канал (цепь) исходного эталонного сигнала
VQEG	Video Quality Experts Group	Экспертная группа по качеству видеоизображений
YUV	Colour Space and file format	Цветовое пространство и формат файла

5 Условные обозначения

Отсутствуют.

6 Описание методов измерения при наличии эталона с ухудшенными характеристиками

6.1 Введение

Хотя отношение PSNR было широко использовано в качестве объективного показателя оценки качества видеоизображений, имеются также сведения, что оно недостаточно хорошо представляет воспринимаемое качество видеоизображения. При анализе того, как люди воспринимают качество видеоизображений, было отмечено, что зрительная система человека чувствительна к ухудшениям по краям. Другими словами, когда краевые пиксели видеоизображения размыты, эксперты, как правило, дают низкие оценки качества видеоизображения даже при высоких отношениях PSNR. Основываясь на этом наблюдении, были разработаны модели, использующие эталон с ухудшенными характеристиками, которые, в основном, измеряют краевые искажения.

На рис. 1 показано, как работает модель с ухудшенным эталоном. Показатели, которые будут использоваться для измерения качества видеоизображения в точке контроля, берутся из исходной видеопоследовательности и передаются. В таблице 1 показаны значения пропускной способности бокового канала для показателей, которые были протестированы в испытании ТВЧ VQEG.

РИСУНОК 1

Блок-схема модели с ухудшенным эталоном

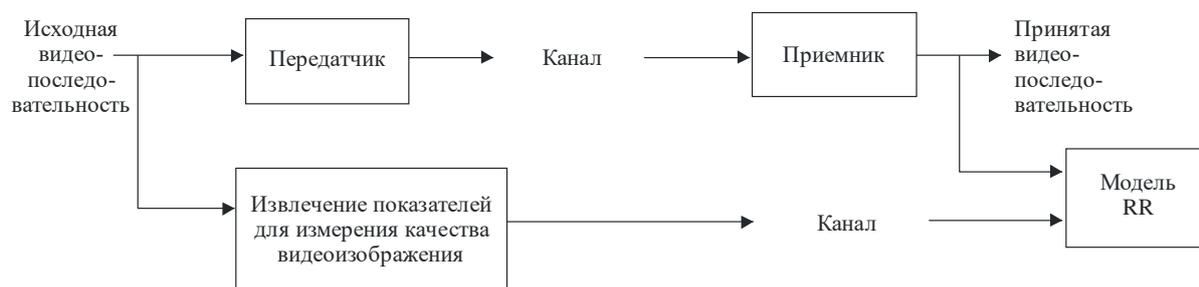


ТАБЛИЦА 1

Пропускная способность бокового канала

Видеоформат	Проверяемые значения пропускной способности
1 080/60 Гц (29,97 к/с) 1 080/30Pp (29,97 к/с)	56 кбит/с, 128 кбит/с, 256 кбит/с
1 080/25Pp (25 к/с) 1 080/50I Гц (25 к/с))	56 кбит/с, 128 кбит/с, 256 кбит/с

6.2 Модель EPSNR с ухудшенным эталоном

6.2.1 Значение краевого отношения PSNR (EPSNR)

Модели с ухудшенным эталоном (RR) главным образом измеряют ухудшения на краях. В этих моделях сначала применяется алгоритм определения краев в исходной видеопоследовательности, чтобы локализовать краевые пиксели. Затем измеряются ухудшения этих краевых пикселей путем расчета среднеквадратической ошибки. Из этой среднеквадратической ошибки вычисляется краевое отношение PSNR.

Можно использовать любой алгоритм определения краев, хотя могут быть незначительные различия в результатах. Чтобы найти краевые пиксели, можно, например, использовать любой оператор градиента. Были предложены несколько операторов градиента. Во многих алгоритмах обнаружения краев сначала вычисляются изображение горизонтального градиента $g_{horizontal}(m,n)$ и изображение вертикального градиента $g_{vertical}(m,n)$ с использованием операторов градиента. Затем изображение амплитуды градиента $g(m,n)$ может быть вычислено следующим образом:

$$g(m,n) = |g_{horizontal}(m,n)| + |g_{vertical}(m,n)|.$$

Наконец, чтобы найти краевые пиксели, к изображениям амплитуды градиента применяется пороговая операция. Другими словами, пиксели, у которых амплитуды градиента превышают пороговое значение, считаются краевыми пикселями.

Эта процедура иллюстрируется на рис. 2–6. На рис. 2 показано исходное изображение. На рис. 3 показано изображение горизонтального градиента $g_{horizontal}(m,n)$, которое получено за счет применения оператора горизонтального градиента к исходному изображению на рис. 2. На рис. 4 показано изображение вертикального градиента $g_{vertical}(m,n)$, которое получено за счет применения оператора вертикального градиента к исходному изображению на рис. 2. На рис. 5 показано изображение амплитуды градиента (изображение края), а на рис. 6 показано двоичное изображение края (изображение маски), полученное путем применения пороговой операции к изображению амплитуды градиента на рис. 5.

РИСУНОК 2

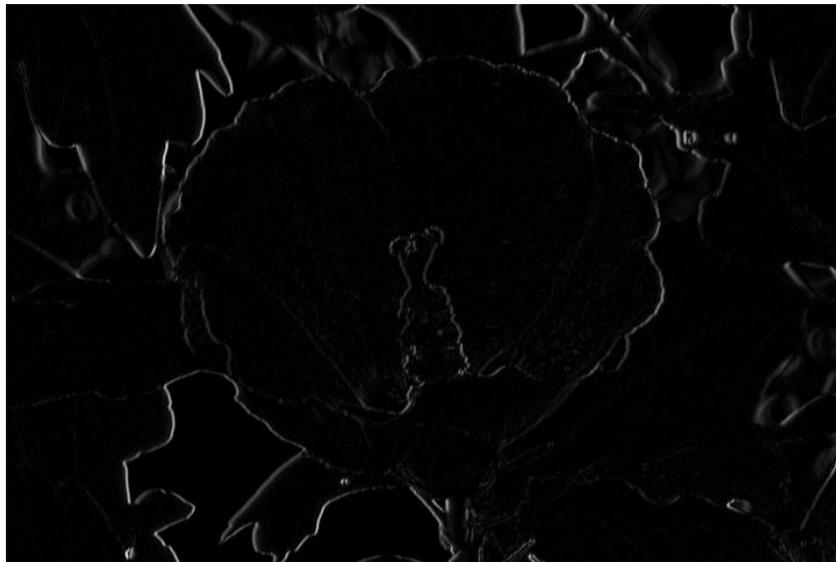
Исходное изображение (оригинальное изображение)



ВТ.1908-02

РИСУНОК 3

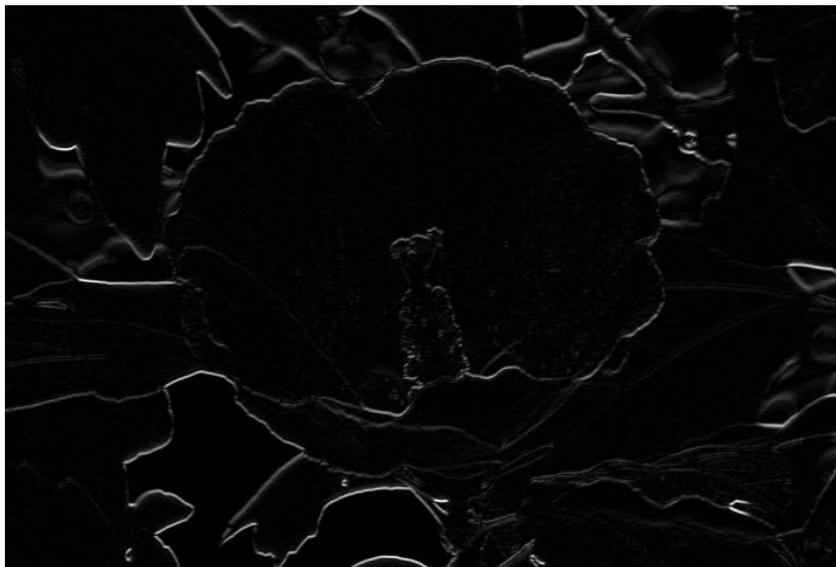
Изображение горизонтального градиента, которое получено за счет применения оператора горизонтального градиента к исходному изображению на рис. 2



ВТ.1908-03

РИСУНОК 4

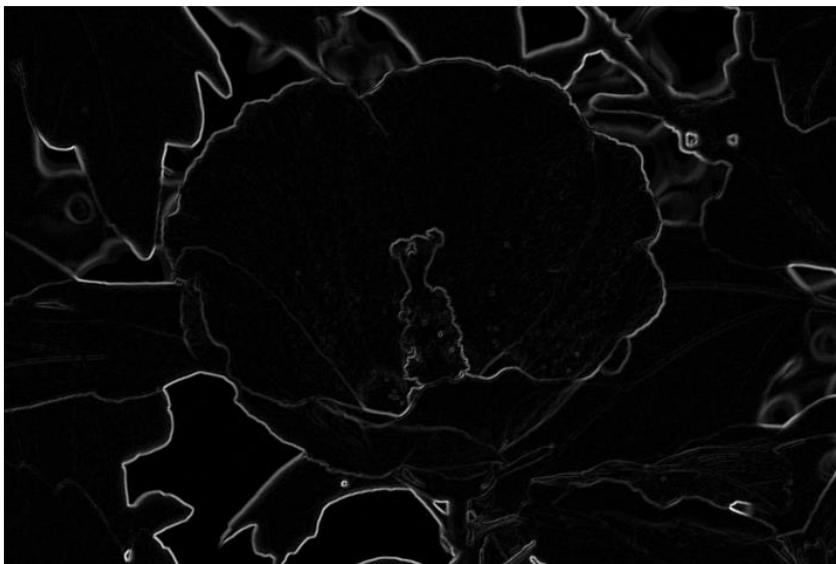
Изображение вертикального градиента, которое получено за счет применения оператора вертикального градиента к исходному изображению на рис. 2



ВТ.1908-04

РИСУНОК 5

Изображение амплитуды градиента



ВТ.1908-05

РИСУНОК 6

Двоичное изображение края (изображение маски),
полученное путем применения пороговой операции
к изображению амплитуды градиента на рис. 5

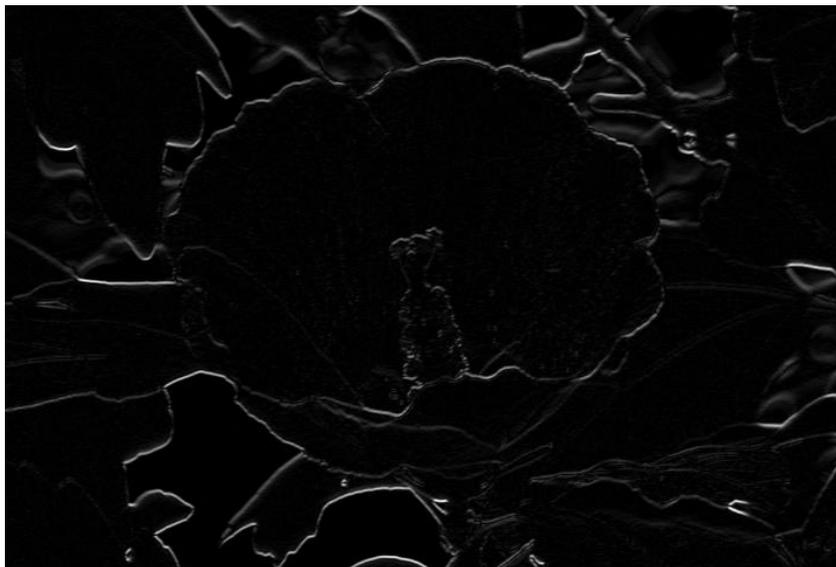


ВТ.1908-06

В качестве варианта, чтобы найти краевые пиксели, можно использовать модифицированную процедуру. Например, можно сначала применить оператор вертикального градиента к исходному изображению, создавая изображение вертикального градиента. Затем для изображения вертикального градиента применяется оператор горизонтального градиента, создавая измененное последующее изображение градиента (изображение горизонтального и вертикального градиентов). Наконец, чтобы найти краевые пиксели, к измененному последующему изображению градиента может быть применена пороговая операция. Другими словами, пиксели измененного последующего изображения градиента, которые превышают пороговое значение, рассматриваются как краевые пиксели. На рис. 7–9 иллюстрируется модифицированная процедура. На рис. 7 показано изображение вертикального градиента $g_{vertical}(m,n)$, которое получено за счет применения оператора вертикального градиента к исходному изображению на рис. 2. На рис. 8 показано изображение измененных последующих градиентов (изображение горизонтального и вертикального градиентов), которое получено за счет использования оператора горизонтального градиента к изображению вертикального градиента на рис. 7. На рис. 9 показано двоичное изображение края (изображение маски), полученное посредством применения пороговой операции к измененному последующему изображению градиента на рис. 8.

РИСУНОК 7

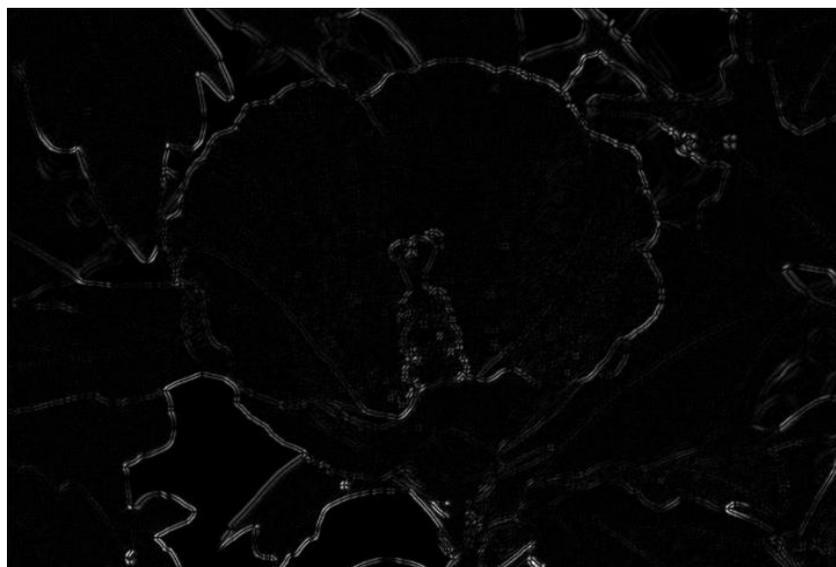
Изображение вертикального градиента, которое получено за счет применения оператора вертикального градиента к исходному изображению на рис. 2



ВТ.1908-07

РИСУНОК 8

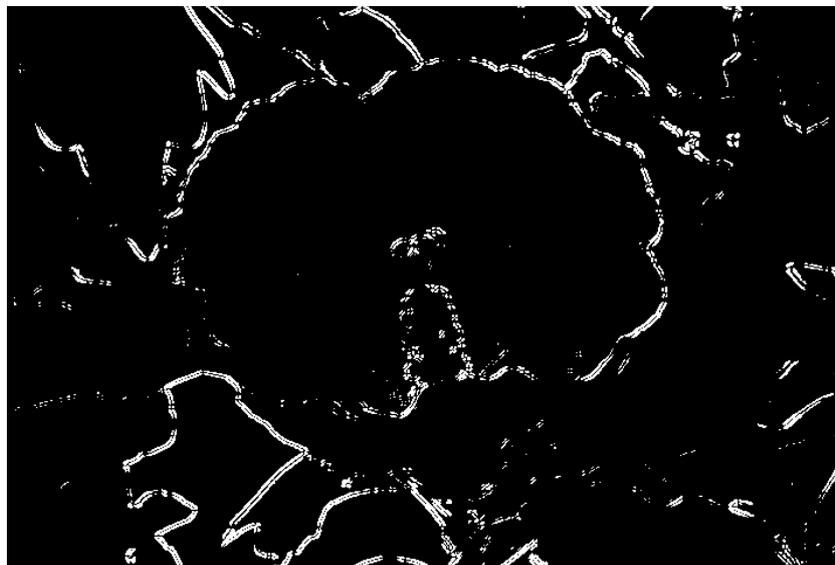
Изображение измененного последующего градиента (изображение горизонтального и вертикального градиентов), которое получено за счет применения оператора горизонтального градиента к изображению вертикального градиента на рис. 7



ВТ.1908-08

РИСУНОК 9

Двоичное изображение края (изображение маски),
полученное за счет применения пороговой операции
к изображению измененного последующего градиента (рис. 8)



ВТ.1908-09

Следует отметить, что оба метода можно рассматривать как алгоритмы определения краев. Можно выбрать любой алгоритм определения краев, в зависимости от характера видеоизображений и алгоритмов сжатия. Тем не менее некоторые методы могут превосходить другие методы.

Таким образом, в модели сначала применяется оператор обнаружения края, в результате чего получаются изображения краев (рис. 5 и 8). Затем создается изображение маски (двоичное изображение краев) путем применения пороговой операции к изображению края (рис. 6 и 9). Другими словами, пиксели изображения края, значения которых меньше, чем порог t_e , устанавливаются в ноль, а пиксели, значения которых равны или превышают порог, устанавливаются в ненулевое значение. На рис. 6 и 9 показаны некоторые изображения масок. Поскольку видеоизображения можно рассматривать как последовательность кадров или полей, вышеуказанные процедуры могут быть применены к каждому кадру или полю видеоизображений. Так как модель может быть использована для видеоизображений на основе полей или видеоизображений на основе кадров, термин "изображение" будет использоваться для обозначения поля или кадра.

6.2.2 Выбор показателей из исходных видеопоследовательностей

Поскольку данная модель является моделью с ухудшенным эталоном (RR), из каждого изображения исходной видеопоследовательности должен быть извлечен тот или иной набор характерных элементов. В модели EPSNR RR из каждого изображения выбирается определенное количество краевых пикселей. Затем кодируются и передаются местоположения и значения пикселей. Однако для некоторых видеопоследовательностей количество краевых пикселей может быть очень малым в случае использования фиксированного порогового значения. В худшем случае оно может быть равным нулю (пустые изображения или изображения с очень низкой частотой). Для разрешения этой проблемы в ситуации, когда число краевых пикселей изображения меньше заданного значения, пользователь может снижать пороговое значение до тех пор, пока количество краевых пикселей не станет больше заданного значения. Кроме того, можно выбрать краевые пиксели, которые соответствуют изображению наибольших значений горизонтального и вертикального градиента. При отсутствии краевых пикселей в кадре (например, в случае пустых изображений) можно произвольно выбрать необходимое количество пикселей или пропустить кадр. Например, если из каждого кадра должны быть выбраны десять краевых пикселей, можно отсортировать пиксели изображений горизонтального и вертикального градиента в зависимости от их значений и выбрать десять максимальных значений. Однако эта процедура может привести к наличию нескольких краевых пикселей в одинаковых местах.

Для разрешения этой проблемы можно сначала выбрать несколько раз нужное количество пикселей из изображения горизонтального и вертикального градиента, а затем случайно выбрать нужное количество краевых пикселей из выбранных пикселей изображений горизонтального и вертикального градиентов. В моделях, испытанных экспертной группой VQEG по ТВЧ, необходимое количество краевых пикселей выбирается случайным образом среди большого пула краевых пикселей. Пул краевых пикселей создается путем применения пороговой операции к изображению градиента.

В моделях EPSNR RR местоположение и значения краевых пикселей кодируются после того, как к выбранным местоположениям пикселей применена фильтрация гауссовским фильтром нижних частот (ФНЧ). Несмотря на то что в испытании VQEG ТВЧ был использован гауссовский ФНЧ (7×3), в зависимости от форматов видеоизображения могут быть использованы различные фильтры нижних частот. Следует отметить, что во время процесса кодирования может быть применена обрезка. Для того чтобы избежать выбора краевых пикселей в обрезанных областях, данная модель выбирает краевые пиксели в средней части (рис. 10). В таблице 2 приведены размеры после обрезки. Таблица 2 также показывает количество битов, необходимых для кодирования местоположений и значений пикселей из краевых пикселей.

ТАБЛИЦА 2

Количество битов, требуемых на один краевой пиксель

Видеоформат	Размер	Размер после обрезки	Биты для определения местоположения	Биты для определения значения пикселя	Всего битов на пиксель
Прогрессивное ТВЧ	1 920 × 1 080	1 856 × 1 032	21	8	29
Чересстрочное ТВЧ	1 920 × 540	1 856 × 516	20	8	28

РИСУНОК 10

Пример обрезки и средней области



ВТ.1908-10

Данная модель выбирает краевые пиксели из каждого кадра в соответствии с разрешенным значением пропускной способности (таблица 1). Таблица 3 показывает количество краевых пикселей на кадр, которое может быть передано для проверяемых значений пропускной способности.

ТАБЛИЦА 3

Количество краевых пикселей на один кадр/поле

Видеоформат	56 кбит/с	128 кбит/с	256 кбит/с
Прогрессивное ТВЧ	46	105	211
Чересстрочное ТВЧ	24	54	109

РИСУНОК 11
Блок-схема модели



ВТ.1908-11

6.2.3 Пространственная/временная регистрация и настройка усиления/смещения

Перед определением разницы между краевыми пикселями исходной видеопоследовательности и обработанной видеопоследовательности, которая является видеопоследовательностью, принятой приемником, данная модель сначала выполняет пространственную/временную регистрацию и настройку усиления/смещения. Использовался метод калибровки, описанный в Рекомендации МСЭ-Т J.244 (Приложение В). Для передачи характеристик усиления и смещения, указанных в Рекомендации МСЭ-Т J.244 (Приложение В), в испытании ТВЧ VQEG были использованы 30% из имеющейся пропускной способности. В случае видеопоследовательности с чересстрочной разверткой метод калибровки применяется три раза: для четных полей, для нечетных полей и для комбинированных кадров, в то время как для прогрессивных видеопоследовательностей метод калибровки применяется для кадров. Если при заданном PSNR разница между ошибкой для четного поля и ошибкой для нечетного поля больше порогового значения, то использовались результаты регистрации (x-сдвиг, y-сдвиг) для меньшего PSNR. В противном случае использовались результаты регистрации для комбинированных кадров. В испытании ТВЧ VQEG порог был установлен равным 2 дБ.

В точке контроля обработанная видеопоследовательность должна быть согласована с краевыми пикселями, извлеченными из исходной видеопоследовательности. Однако, если пропускная способность бокового канала мала, доступны всего несколько краевых пикселей исходной видеопоследовательности (рис. 12). Таким образом, временная регистрация может быть неточной, если она осуществляется с помощью одного кадра (рис. 13). Для решения этой проблемы в модели используется окно для временной регистрации. Вместо использования одного кадра обработанной видеопоследовательности, модель создает окно, которое состоит из нескольких соседних кадров, чтобы найти оптимальный временной сдвиг. На рис. 14 показана эта процедура. Среднеквадратическая ошибка в пределах окна вычисляется следующим образом:

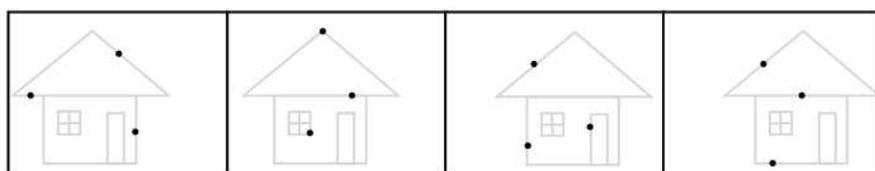
$$MSE_{window} = \frac{1}{N_{win}} \sum (E_{SRC}(i) - E_{PVS}(i))^2,$$

где: MSE_{window} – это среднеквадратическая ошибка в пределах окна, $E_{SRC}(i)$ – крайовой пиксель в пределах окна, который имеет соответствующий пиксель в обработанной видеопоследовательности, $E_{PVS}(i)$ – пиксель обработанной видеопоследовательности, соответствующий крайовому пикселю, и N_{win} – общее количество используемых крайевых пикселей для вычисления MSE_{window} . Эта среднеквадратическая ошибка в пределах окна используется в качестве разницы между кадром обработанной видеопоследовательности и соответствующим кадром исходной последовательности.

Размер окна может быть определен с учетом характера обработанной видеопоследовательности. Для типичного применения рекомендуется окно, соответствующее двум секундам передачи. В иных случаях могут быть применены различные размеры окон и выбирается наилучший, который обеспечивает наименьшую среднеквадратичную ошибку. Кроме того, различные центры окон могут быть использованы для рассмотрения пропуска кадров из-за ошибок при передаче данных (рис. 18).

РИСУНОК 12

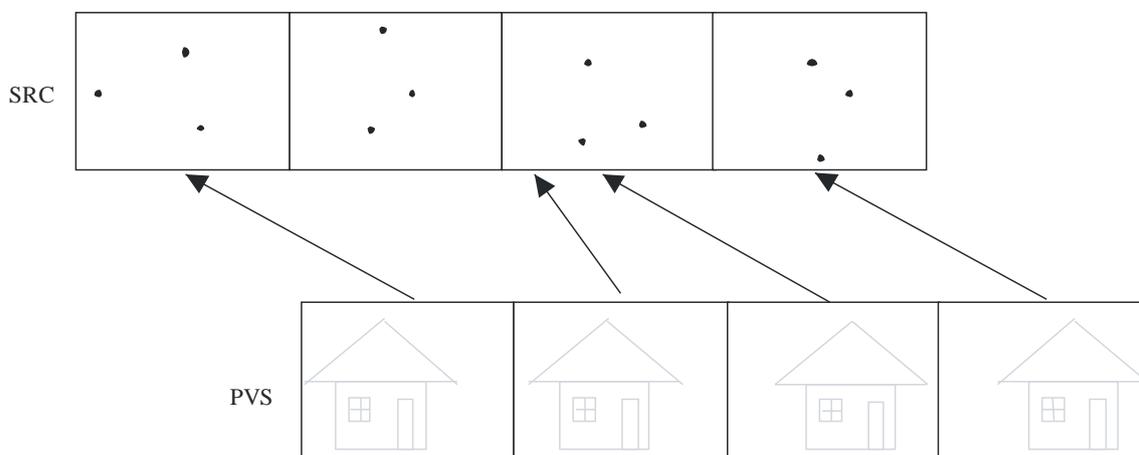
Выбор крайевых пикселей исходной видеопоследовательности



ВТ.1908-12

РИСУНОК 13

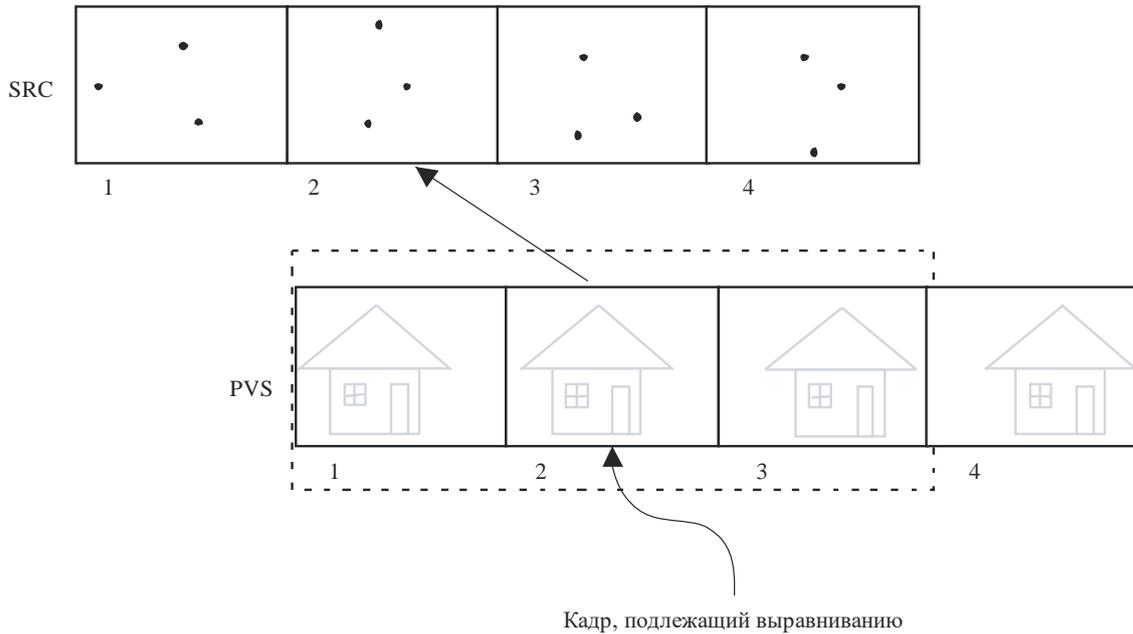
Выравнивание обработанной видеопоследовательности по крайевым пикселям исходной видеопоследовательности



ВТ.1908-13

РИСУНОК 14

Выравнивание обработанной видеопоследовательности по краевым пикселям, с использованием окна

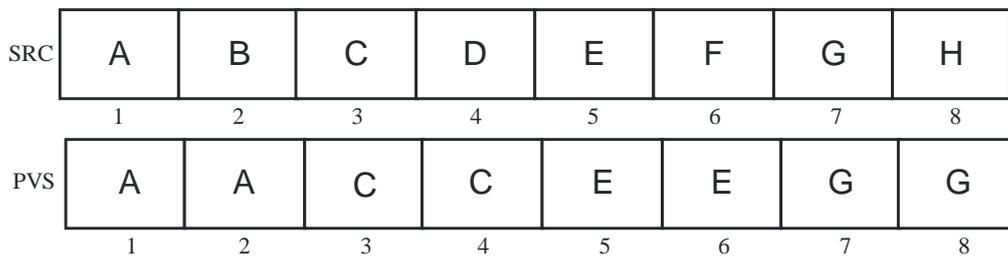


ВТ.1908-14

Когда исходная видеопоследовательность кодируется с высокой степенью сжатия, кодер может уменьшить количество кадров в секунду, а обработанная видеопоследовательность имеет повторяющиеся кадры (рис. 15). На рис. 15 обработанная видеопоследовательность не имеет кадров, соответствующих некоторым кадрам исходной видеопоследовательности (2, 4, 6, 8-й кадры). В этом случае модель не использует повторяющиеся кадры при вычислении среднеквадратической ошибки. Другими словами, модель выполняет временную регистрацию, используя первый кадр (достоверный кадр) каждого повторяющегося блока. Таким образом, на рис. 16 всего три кадра (3, 5, 7-й кадры) в пределах окна используются для временной регистрации.

РИСУНОК 15

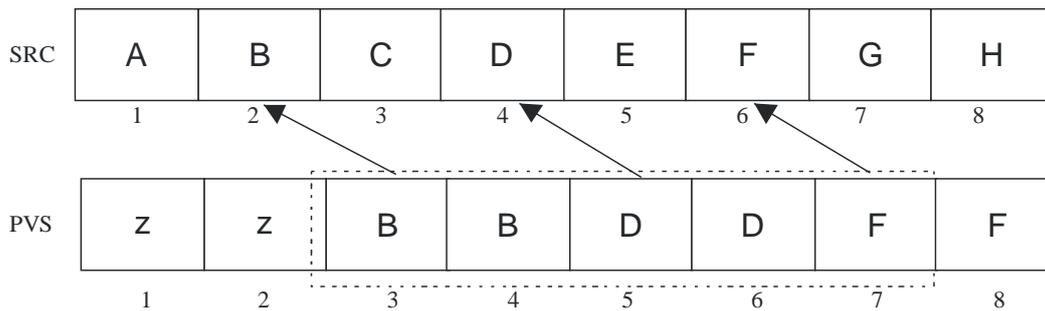
Пример повторяющихся кадров



ВТ.1908-15

РИСУНОК 16

Обработка повторяющихся кадров

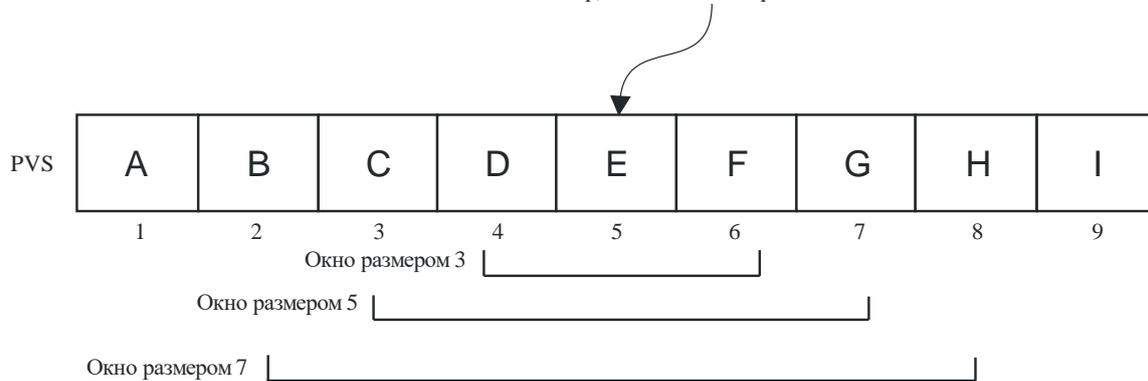


ВТ.1908-16

РИСУНОК 17

Окна равных размеров

Кадр, подлежащий выравниванию

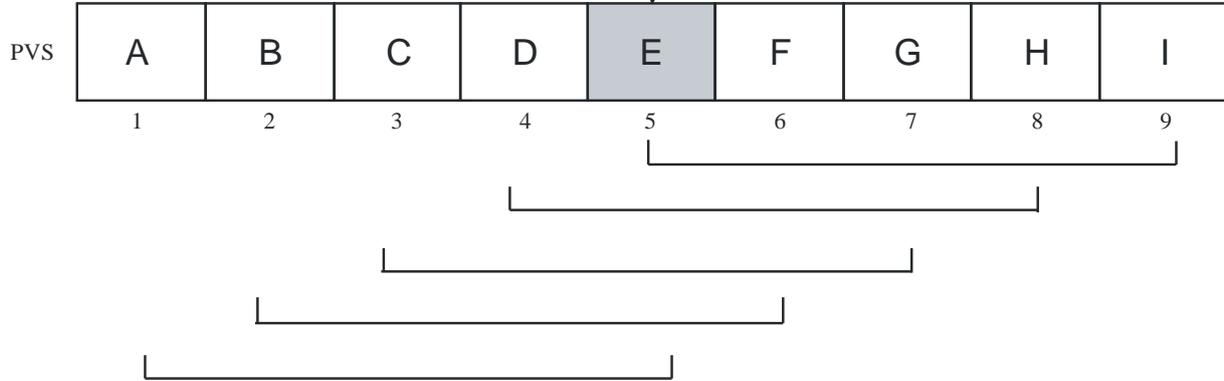


ВТ.1908-17

РИСУНОК 18

Центры окон

Кадр, подлежащий выравниванию



ВТ.1908-18

6.2.4 Вычисление EPSNR и постобработка

После выполнения временной регистрации, вычисляется среднее значение разницы между краевыми пикселями исходной видеопоследовательности и соответствующими пикселями обработанной видеопоследовательности, которое может быть определено как краевая среднеквадратическая ошибка обработанной видеопоследовательности (MSE_{edge}). Окончательно значение EPSNR (краевое PSNR) вычисляется следующим образом:

$$EPSNR = 10 \log_{10} \left(\frac{P^2}{MSE_{edge}} \right),$$

где p – пиковое значение изображения.

Поскольку различные нарушения могут снизить качество видеоизображения, значение EPSNR корректируется с учетом этих эффектов, которые количественно определяются в следующих подразделах.

1) Показатель блочности I

Для рассмотрения эффектов блочности вычисляются средние значения разницы в колонке. Показатель блочности для i -го кадра вычисляется по модулю 8 следующим образом:

$$Blk[i] = \frac{\text{наибольшая разница в колонке}}{\text{вторая по величине разница в колонке}}.$$

Окончательный показатель блочности (*Blocking*) вычисляется путем усреднения показателей блочности кадров.

$$\text{Блочность} = \frac{1}{\text{количество кадров}} \sum_i Blk[i].$$

Окончательно используются следующие уравнения:

$$\begin{aligned} \text{IF}(\text{BLOCKING} > 12 \text{ and } 25 \leq \text{EPSNR} < 30) \quad \text{adjust_EPSNR_blk1} &= 3 \\ \text{IF}(\text{BLOCKING} > 5 \text{ and } 30 \leq \text{EPSNR} < 35) \quad \text{adjust_EPSNR_blk1} &= 5 \end{aligned}$$

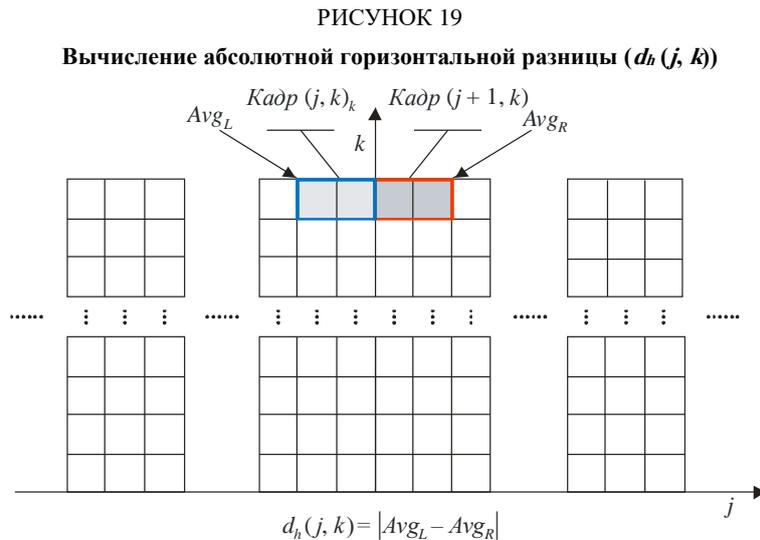
2) Показатель блочности II

Если предположить, что ухудшения из-за блочности могут возникать в каждом 8-м столбце (например, в MPEG2), то также используется второй показатель блочности. Для вычисления второго показателя блочности сначала вычисляется абсолютная разница по горизонтали следующим образом (рис. 19):

$$d_h(j, k) = |Avg_L - Avg_R|,$$

где:

$$Avg_L = \frac{1}{2} \sum_{p=-1}^0 Frame(j+p, k), \quad Avg_R = \frac{1}{2} \sum_{p=1}^2 Frame(j+p, k).$$



ВТ.1908-19

Затем определяется сумма блочности по горизонтали (SB_h) в позиции j следующим образом:

$$SB_h[j] = \left(\sum_{1 \leq k \leq height} (|Frame(j, k) - Frame(j+1, k)| \times u(d_h(j, k) - \Phi(Avg_L))) \right)^2,$$

где $u(\cdot)$ представляет единичную ступенчатую функцию и:

$$\Phi(s) = \begin{cases} 17(1 - \sqrt{s/127}) + 3 & \text{при } s \leq 127 \\ 3(s - 127)/128 + 3 & \text{другие} \end{cases}.$$

После повторения этой процедуры для целых кадров кадровая блочность по горизонтали (FB_h) вычисляется следующим образом:

$$FB_h = \left(\sum_{\substack{1 \leq j \leq width \\ j \equiv 0 \pmod{8}}} SB_h(j) \right)^{1/2}.$$

Разница по столбцам (NFB_h) для каждого кадра, исключая каждый 8-й столбец, вычисляется следующим образом:

$$NFB_h = \frac{1}{7} \sum_{l=1}^7 \left(\sum_{\substack{1 \leq j \leq width \\ j=l \pmod{8}}} \left(\sum_{1 \leq k \leq height} (|Frame(j,k) - Frame(j+1,k)| \times u(d_h(j,k) - \Phi(Avg_L))) \right)^2 \right)^{1/2}.$$

Затем окончательный показатель блочности по горизонтали, BLK_H , вычисляется следующим образом:

$$BLK_H = \ln(FB_h / NFB_h).$$

Показатель блочности по вертикали, BLK_V , вычислялся аналогично. Для чересстрочных видеопоследовательностей показатель блочности по вертикали вычисляется из последовательности полей. Оценка блочности i -го кадра вычисляется следующим образом:

$$FrameBLK(i) = 0,5 \times BLK_H + 0,5 \times BLK_V.$$

Окончательный показатель блочности ($BLOCKING2$) вычисляется путем усреднения верхних 10% показателей блочности кадров.

Окончательно, используются следующие уравнения:

IF(BLOCKING2 > 1.5 and 25 ≤ EPSNR<30)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1.3 and 30 ≤ EPSNR<35)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1.5 and 35 ≤ EPSNR<40)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1 and 40 ≤ EPSNR<45)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 0.5 and 45 ≤ EPSNR<55)	adjust_EPSNR_blk2=2

Как видно из приведенных выше уравнений, это выравнивание имеет незначительное влияние на окончательное значение EPSNR. Если артефакты блочности не встречаются в каждом 8-м столбце, можно пропустить эту процедуру выравнивания или сначала найти местоположения артефактов блочности. Кроме того, можно использовать различные функции для $\Phi(s)$.

3) Максимальное количество статических кадров и общее количество статических кадров

Ошибки при передаче могут привести к длинной последовательности статических кадров. Для оценки таких статических кадров используются следующие уравнения:

IF(MAX_FREEZE ≥ 8 and 25 ≤ EPSNR<30)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 6 and 30 ≤ EPSNR<35)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 3 and 35 ≤ EPSNR<40)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 1.5 and 40 ≤ EPSNR<45)	adjust_EPSNR_max_freeze=2
IF(MAX_FREEZE ≥ 1 and 45 ≤ EPSNR<95)	adjust_EPSNR_max_freeze=2

где MAX_FREEZE – это наибольшая продолжительность статических кадров. Следует отметить, что если видеопоследовательность отсутствует в течение 10 секунд, должны быть использованы другие пороговые значения.

Также общее количество статических кадров определяется следующим образом:

IF(TOTAL_FREEZE ≥ 80 and 25 ≤ EPSNR<30)	adjust_EPSNR_total_freeze=3
IF(TOTAL_FREEZE ≥ 40 and 30 ≤ EPSNR<35)	adjust_EPSNR_total_freeze=4
IF(TOTAL_FREEZE ≥ 10 and 35 ≤ EPSNR<40)	adjust_EPSNR_total_freeze=3.5
IF(TOTAL_FREEZE ≥ 2 and EPSNR ≥ 40)	adjust_EPSNR_total_freeze=1.5

где TOTAL_FREEZE – это общая продолжительность статических кадров. Следует отметить, что если видеопоследовательность отсутствует в течение 10 секунд, должны быть использованы другие пороговые значения.

4) Блок ошибок передачи

Локальные блоки статических кадров могут иметь место из-за ошибок при передаче. Кроме того, в статичных сценах некоторые блоки идентичны блокам предыдущих кадров на тех же позициях. Для рассмотрения локальных блоков статических кадров из-за ошибок передачи, блоки, которые содержат переданные краевые пиксели, классифицируются либо как идентичные блоки (т. е. блоки идентичны блокам предыдущих кадров), либо как различные блоки. Затем вычисляются два отношения EPSNR для идентичных блоков и различных блоков. Если разница между двумя отношениями EPSNR (EPSNR_diff) велика, это означает, что при передаче могли возникнуть ошибки. Основываясь на этом наблюдении, значение EPSNR корректируется следующим образом:

IF($8 \leq \text{EPSNR_diff} \leq 30$ and $25 \leq \text{EPSNR} < 30$)	adjust_EPSNR_diff= 3
IF($9 \leq \text{EPSNR_diff} \leq 30$ and $30 \leq \text{EPSNR} < 35$)	adjust_EPSNR_diff= 4
IF($10 \leq \text{EPSNR_diff} \leq 30$ and $35 \leq \text{EPSNR} < 40$)	adjust_EPSNR_diff= 6
IF($9 \leq \text{EPSNR_diff} < 10$ and $35 \leq \text{EPSNR} < 40$)	adjust_EPSNR_diff= 2
IF($9 \leq \text{EPSNR_diff} \leq 30$ and $40 \leq \text{EPSNR} < 45$)	adjust_EPSNR_diff= 4

Однако, если общее количество идентичных блоков меньше 100, корректировка не производится.

5) Окончательная корректировка значения EPSNR

Окончательно значение EPSNR корректируется следующим образом:

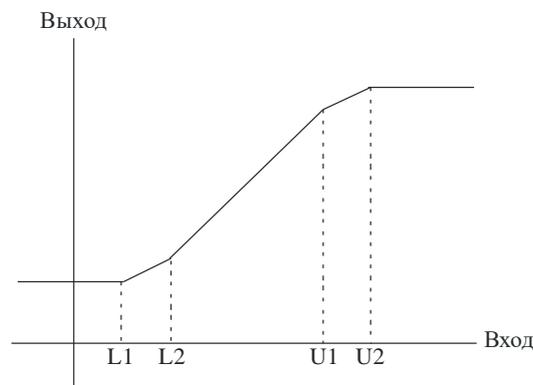
$$\text{EPSNR} \leq \text{EPSNR} - \text{MAX}(\text{adjust_EPSNR_blk1}, \text{adjust_EPSNR_blk2}, \text{adjust_EPSNR_max_freeze}, \text{adjust_EPSNR_total_freeze}, \text{adjust_EPSNR_diff})$$

6) Кусочно-линейная аппроксимация

Когда EPSNR превышает некоторое конкретное значение, воспринимаемое качество достигает предельного уровня. В этом случае можно установить верхнюю границу отношения EPSNR. Кроме того, если желательна линейная зависимость между EPSNR и DMOS (разностный показатель усредненного мнения), возможно применение кусочно-линейной функции, как показано на рис. 20. В тестируемой модели для ТВЧ VQEG верхняя граница была установлена на 50, а нижняя граница – на 19.

РИСУНОК 20

Кусочно-линейная функция для линейной зависимости между EPSNR и DMOS



Дополнение 1 (информационное)

Выводы экспертной группы по качеству видеоизображений (VQEG)

Исследования измерений воспринимаемого качества видеоизображения проводятся в рамках неофициальной группы, называемой экспертная группа по качеству видеоизображений (VQEG), которая отчитывается перед 9-й и 12-й Исследовательскими комиссиями МСЭ-Т. Недавно завершенная первая фаза тестирования телевидения высокой четкости, проведенная VQEG, оценила эффективность предлагаемых алгоритмов измерения воспринимаемого качества видеоизображения при наличии полного эталона.

Нижеприведенные статистические данные взяты из заключительного доклада VQEG по ТВЧ. Следует отметить, что текст доклада VQEG по ТВЧ включает другие показатели, в частности корреляцию Пирсона и среднеквадратическую ошибку (RMSE), рассчитанные для отдельных экспериментов, доверительные интервалы, проверку на статистическую значимость в отдельных экспериментах, анализ подмножеств данных, которые включают специфические искажения (например, только при кодировании согласно H.264 ITU-T), диаграммы рассеяния, а также коэффициенты соответствия.

Первичный анализ

Характеристики модели RR представлены в таблице 4. Отношение PSNR вычисляется в соответствии с Рекомендацией МСЭ-Т J.340 и включено в этот анализ в целях сравнения. Графа "RMSE супермножества" определяет основной показатель (RMSE), рассчитанный на агрегированном супермножестве (т. е. все шесть экспериментов, отображенные на одной шкале). Графа "Наилучший общий показатель в группе" определяет количество экспериментов (от 0 до 6), для которых эта модель была либо наиболее эффективной моделью, либо статистически эквивалентна наиболее эффективной модели. Графа "Лучше, чем общее PSNR" определяет количество экспериментов (от 0 до 6), для которых модель была статистически лучше, чем PSNR. Графа "Лучше, чем PSNR супермножества" содержит список того, является ли каждая модель статистически лучшей, чем PSNR на агрегированном супермножестве. Графа "Корреляция супермножества" определяет корреляцию Пирсона, рассчитанную на агрегированном супермножестве.

ТАБЛИЦА 4

Показатель	PSNR	Yonsei56k	Yonsei128k	Yonsei256k
RMSE супермножества	0,71	0,73	0,73	0,73
Наилучший общий показатель в группе	6	4	4	4
Эквивалентно или лучше, чем общее PSNR	6	4	4	4
Эквивалентно PSNR супермножества	Да	Да	Да	Да
Корреляция супермножества	0,78	0,77	0,77	0,77

Ввиду того что эффективность модели статистически одинакова для всех трех значений пропускной способности, рекомендуется использовать эту модель с пропускной способностью бокового канала, равной по крайней мере 56 кбит/с.

Вторичный анализ

В таблице 5 приведено отношение RMSE для модели RR для подгрупп супермножества. Эти подгруппы разделяют данные по типу кодирования (ITU-T H.264 или MPEG-2), а также по наличию ошибок при передаче данных (Ошибки) или по тому, содержала ли HRC только артефакты кодирования (Кодирование). Поскольку эксперименты не предназначались для того, чтобы эти переменные равномерно охватывали весь спектр качества, только RMSE представлены для этих подгрупп.

ТАБЛИЦА 5

Тип HRC	PSNR	Yonsei56k	Yonsei128k	Yonsei256k
Кодирование H.264	0,75	0,65	0,65	0,65
Ошибка H.264	0,67	0,86	0,85	0,86
Кодирование MPEG-2	0,78	0,81	0,81	0,80
Ошибка MPEG-2	0,66	0,68	0,68	0,68
Кодирование	0,75	0,69	0,69	0,69
Ошибка	0,67	0,79	0,78	0,79

Библиография

- Recommendation ITU-T P.910 (2008), *Subjective video quality assessment methods for multimedia applications*.
- Recommendation ITU-T P.911 (1998), *Subjective audiovisual quality assessment methods for multimedia applications*.
- Recommendation ITU-T J.143 (2000), *User requirements for objective perceptual video quality measurements in digital cable television*.
- Рекомендации МСЭ-R ВТ.500, *Методика субъективной оценки качества телевизионных изображений*.
- Recommendation ITU-T J.340 (2010), *Reference algorithm for computing peak signal to noise ratio of a processed video sequence with compensation for constant spatial shifts, constant temporal shift, and constant luminance gain and offset*.
-