

**Рекомендация МСЭ-R ВТ.1885**  
(03/2011)

**Методы объективного измерения  
воспринимаемого качества изображения,  
предназначенные для цифрового  
вещательного телевидения стандартной  
четкости, при наличии эталонного  
сигнала с уменьшенной полосой частот**

**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-T/МСЭ-R/ИСО/МЭК, упоминаемой в Приложении 1 к Резолюции 1 МСЭ-R. 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-T/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)
<b>BT</b>	<b>Радиовещательная служба (телевизионная)</b>
F	Фиксированная служба
M	Подвижная спутниковая служба, спутниковая служба радиоопределения, любительская спутниковая 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 1 МСЭ-R.

Электронная публикация  
Женева, 2011 г.

© ITU 2011

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.1885

**Методы объективного измерения воспринимаемого качества изображения, предназначенные для цифрового вещательного телевидения стандартной четкости, при наличии эталонного сигнала с уменьшенной полосой частот**

(2011)

**Сфера применения**

В настоящей Рекомендации описаны методы объективной оценки качества изображения, предназначенные для цифрового вещательного телевидения стандартной четкости, которые позволяют измерять воспринимаемое качество изображения в условиях приема на мобильные и стационарные средства, когда характеристики, полученные на основании эталонного видеосигнала, легкодоступны в измерительной точке.

Ассамблея радиосвязи МСЭ,

*учитывая,*

- a) что желательно обеспечивать способность автоматического измерения ухудшения транслируемых видеосигналов;
- b) что воспринимаемое качество изображения при приеме на мобильные средства может динамически меняться в зависимости от условий приема;
- c) что объективное измерение воспринимаемого качества изображения может эффективно дополнять методы субъективной оценки;
- d) что на рассмотрение МСЭ-R были предложены три метода измерения качества изображения для цифрового вещательного телевидения стандартной четкости при наличии эталонного сигнала с уменьшенной полосой частот, и было обнаружено, что эти методы обеспечивают эквивалентные и хорошо согласованные результаты;
- e) что в Рекомендации МСЭ-R ВТ.1683 описываются объективные методы измерения воспринимаемого качества изображения цифрового радиовещательного телевидения стандартной четкости при наличии эталонного сигнала с полной полосой частот,

*рекомендует,*

**1** чтобы для объективного измерения воспринимаемого качества изображения для цифрового радиовещательного телевидения стандартной четкости при наличии эталонного сигнала с уменьшенной полосой частот использовалась любая из приведенных в Приложении 1 моделей объективной оценки качества изображения.

**1 Введение**

Данные испытания RRNR-TV предназначены для изображений, определенных в Рекомендации МСЭ-R ВТ.601-6, и двух типов моделей с эталонным сигналом с уменьшенной полосой частот (ухудшенным эталонным сигналом) (RR) и отсутствием эталонного сигнала (NR). Модели RR имеют доступ с ограниченной шириной полосы к видеосигналу источника, а модели NR не имеют доступа к видеосигналу источника.

В цепях HRC в каждом эксперименте осуществлялось и кодирование только артефактов, и кодирование с ошибками передачи. Были проверены схемы кодирования MPEG-2 и H.264 (MPEG-4 часть 10). Кодеры MPEG-2 работали с разной скоростью передачи в диапазоне от 1,0 до 5,5 Мбит/с. Кодеры H.264 работали с разной скоростью передачи в диапазоне от 1,0 до 3,98 Мбит/с. Каждый эксперимент включал 12 исходных последовательностей, при этом два источника были секретными. В каждом эксперименте применялись 34 HRC и 156 обработанных видеопоследовательностей (PVS), из которых 40 PVS содержали ошибки передачи, а 116 – только ошибки кодирования.

## 1.1 Применение

В настоящей Рекомендации представлены оценки качества изображения для изображений классов от TV3 до MM5B, которые определены в Приложении В к Рекомендации МСЭ-Т Р.911. Применение моделей оценки, описанных в настоящей Рекомендации, включает в том числе следующее:

- 1 контроль качества, потенциально в режиме реального времени, у источника в процессе эксплуатации;
- 2 контроль качества в удаленной точке, когда доступны боковые каналы для функций, полученных из видеопоследовательностей источника;
- 3 измерение качества для контроля системы хранения или передачи, в которой используются методы сжатия и развертывания видеосигнала при либо однократном прохождении, либо при использовании комплекса таких методов;
- 4 лабораторное испытание видеосистем.

## 1.2 Ограничения

Описанные в данной Рекомендации модели оценки не могут быть использованы для полной замены субъективных испытаний. Значения корреляции между двумя тщательно разработанными и выполненными субъективными испытаниями (в двух разных лабораториях) обычно находятся в диапазоне 0,95–0,98. Если данная Рекомендация используется для сравнения двух разных кодеков для определения точности модели в данных конкретных условиях рекомендуется использовать количественный метод (например, описанный в Рекомендации МСЭ-Т J.149).

Модели, представленные в настоящей Рекомендации, были проверены при помощи измерения объективного качества изображения, демонстрирующего фиксацию кадров до 2 с.

Модели настоящей Рекомендации не проверялись в отношении измерения объективного качества изображения, имеющего постоянно растущую задержку (например, видеосигнала, который не отбрасывает пропущенные кадры после фиксации кадра).

Следует отметить, что в случае новых технологий кодирования и передачи, создающих артефакты изображения, не включенные в данную оценку, модели объективной оценки могут дать неверные результаты. В таком случае необходима субъективная оценка.

## 2 Справочные документы

Указанные ниже Рекомендации МСЭ и другие справочные документы содержат положения, которые путем ссылки на них в данном тексте составляют положения настоящей Рекомендации. На момент публикации указанные издания были действующими. Все Рекомендации и другие справочные документы могут подвергаться пересмотру, поэтому пользователям настоящей Рекомендации предлагается изучить возможность применения последнего издания Рекомендаций и других справочных документов, перечисленных ниже. Список действующих в настоящее время Рекомендаций МСЭ-Т регулярно публикуется.

Ссылка на документ в рамках этой Рекомендации не дает ему как отдельному документу статуса Рекомендации.

### 2.1 Нормативные справочные документы

Рекомендация МСЭ-R ВТ.500-12 – Методика субъективной оценки качества телевизионных изображений.

ITU-T Recommendation P.910 (2008) – Subjective video quality assessment methods for multimedia applications.

ITU-T Recommendation P.911 (1998) – Subjective audiovisual quality assessment methods for multimedia applications.

ITU-T Recommendation J.143 (2000) – User requirements for objective perceptual video quality measurements in digital cable television.

ITU-T Recommendation 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.

## 2.2 Информационные справочные документы

ITU-T Recommendation J.149 (1998) – Subjective audiovisual quality assessment methods for multimedia applications.

ITU-T Recommendation J.144 (2001) – Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference.

ITU-T Recommendation P.931 (1998) – Multimedia communications delay, synchronization and frame rate measurement.

ITU-T Recommendation J.148 (2003) – Requirements for an objective perceptual multimedia quality model.

ITU-T Recommendation H.261 (1993) – Video codec for audiovisual services at p x 64 kbits.

Рекомендация МСЭ-Т H.263 (1996 г.) – Кодирование видеосигнала для низкоскоростной связи.

Рекомендация МСЭ-Т H.263 (1998 г.) – Кодирование видеосигнала для низкоскоростной связи (H.263+).

Рекомендация МСЭ-Т H.264 (2003 г.) – Улучшенное кодирование видеосигнала для основополагающих аудиовизуальных услуг.

VQEG – Validation of reduced-reference and no-reference objective models for standard definition television, Phase I, 2009.

## 3 Определения

### 3.1 Термины, определенные в других документах

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

**3.1.1 субъективная оценка (изображение) (subjective assessment (picture))** (ITU-T Recommendation J.144): факультативное цитируемое определение.

**3.1.2 объективное измерение на основе восприятия (изображение) (objective perceptual measurement (picture))** (ITU-T Recommendation J.144): факультативное цитируемое определение.

**3.1.3 сторонник (Proponent)** (ITU-T Recommendation J.144): факультативное цитируемое определение.

### 3.2 Термины, определенные в настоящей Рекомендации

В настоящей Рекомендации определены следующие термины:

**3.2.1 Анормальный повтор кадров (Anomalous frame repetition):** определяется как событие, при котором на выходе НРС постоянно вырабатывается один кадр в ответ на необычное или экстраординарное событие. Анормальный повтор кадров включает, в том числе, события следующих типов: ошибка в канале передачи, изменение задержки в канале передачи, ограниченные вычислительные ресурсы, воздействующие на производительность декодера, а также ограниченные компьютерные ресурсы, воздействующие на отображение видеосигнала.

**3.2.2 Постоянный пропуск кадра (Constant frame skipping):** определяется как событие, при котором на выходе НРС вырабатываются кадры с обновленным содержимым при эффективной частоте кадров, которая является фиксированной и меньше частоты кадров источника.

**3.2.3 Эффективная частота кадров (Effective frame rate):** определяется как число уникальных кадров (т. е. общее число кадров – повторные кадры) в секунду.

**3.2.4 Частота кадров (Frame rate):** определяется как число уникальных кадров (т. е. общее число кадров – повторные кадры) в секунду.

**3.2.5 Заданная частота кадров (Intended frame rate):** определяется как число видеок кадров в секунду, физически сохраненных для представления видеопоследовательности. Частота кадров должна быть постоянной. Двумя примерами постоянной заданной частоты кадров являются лента BetacamSP®, содержащая 25 кадров в секунду, и YUV-файл на 625 строк, соответствующий VQEG FR-TV Phase I и содержащий 25 кадров в секунду; в обоих случаях заданная частота кадров составляет 25 кадров в секунду.

**3.2.6 Реальные сетевые условия (Live network conditions):** определяются как ошибки, возникающие в цифровом двоичном видеопотоке под воздействием реальных сетевых условий.

**3.2.7 Пауза с пропуском (Pausing with skipping):** определяется как событие, при котором изображение приостанавливается на какой-либо период времени и далее возобновляется с потерей видеoinформации. При возникновении пауз с пропуском временная задержка в системе изменяется относительно среднего значения задержки системы, увеличиваясь или уменьшаясь. Одним из примеров паузы с пропуском является пара IP-видеотелефонов, когда интенсивный сетевой трафик вызывает короткие замирания изображения на IP-видеотелефоне, когда изображение на IP-видеотелефоне возобновляется, часть содержимого оказывается потерянной. Постоянный пропуск кадров и переменный пропуск кадров являются подмножествами паузы с пропуском. Длительность обработанной видеопоследовательности, содержащей паузы с пропусками, будет примерно такой же, как и длительность соответствующей оригинальной видеопоследовательности.

**3.2.8 Пауза без пропуска (Pausing without skipping):** определяется как любое событие, при котором изображение приостанавливается на какой-либо период времени и далее возобновляется без потери видеoinформации. Таким образом, временная задержка в системе должна увеличиваться.

**3.2.9 Частота обновления (Refresh rate):** определяется как частота, с которой обновляется изображение.

**3.2.10 Имитируемые ошибки передачи (Simulated transmission errors):** определяются как ошибки, возникающие в цифровом двоичном видеопотоке в жестко регулируемой среде. К примерам относятся частота имитируемой потери пакетов и имитируемые битовые ошибки.

**3.2.11 Частота кадров источника (Source frame rate) (SFR):** заданная частота кадров оригинальной видеопоследовательности источника. Частота кадров источника является постоянной. Для испытаний VQEG RRNR-TV частота SFR составляла 25 или 30 кадров в секунду.

**3.2.12 Ошибки передачи (Transmission errors):** определяются как любая ошибка, возникающая в ходе передачи видеосигналов. Примерами этого типа ошибок могут служить имитируемые ошибки передачи и реальные сетевые условия.

**3.2.13 Переменный пропуск кадров (Variable frame skipping):** определяется как событие, при котором на выходе HRC вырабатываются кадры с обновленным содержимым при эффективной частоте кадров, которая изменяется во времени. Временная задержка в системе будет с течением времени увеличиваться и уменьшаться, изменяясь относительно среднего значения задержки системы. Длительность обработанной видеопоследовательности, содержащей переменные пропуски кадров, будет примерно такой же, как и длительность соответствующей исходной видеопоследовательности.

#### 4 Сокращения и акронимы

В настоящей Рекомендации используются следующие сокращения и акронимы:

ACR	Absolute category rating (see ITU-T Recommendation P.910)	Показатель абсолютной категории (см. Рекомендацию МСЭ-Т P.910)
ACR-HR	Absolute category rating with hidden reference (see ITU-T Recommendation P.910)	Показатель абсолютной категории со скрытой эталонной версией (см. Рекомендацию МСЭ-Т P.910)
AVI	Audio video interleave	Чередование аудио и видео
DMOS	Difference mean opinion score	Средняя субъективная оценка разницы
FR	Full reference	Эталонный сигнал с полной полосой частот (полный эталонный сигнал)
FRTV	Full reference television	Телевидение с полным эталонным сигналом

HRC	Hypothetical reference circuit	Гипотетическая эталонная цепь
NR	No (or Zero) reference	Отсутствие эталонного сигнала (нулевой эталонный сигнал)
PSNR	Peak signal-to-noise ratio	Пиковое отношение сигнал/шум
PVS	Processed video sequence	Обработанная видеопоследовательность
RMSE	Root mean square error	Среднеквадратичная ошибка
RR	Reduced reference	Эталонный сигнал с уменьшенной полосой частот (ухудшенный эталонный сигнал)
SFR	Source frame rate	Частота кадров источника
SRC	Source reference channel or circuit	Эталонный канал или эталонная цепь источника
VQEG	Video Quality Experts Group	Группа экспертов по качеству видеосигнала
YUV	Colour space	Цветовое пространство

## 5 Условные обозначения

Отсутствуют.

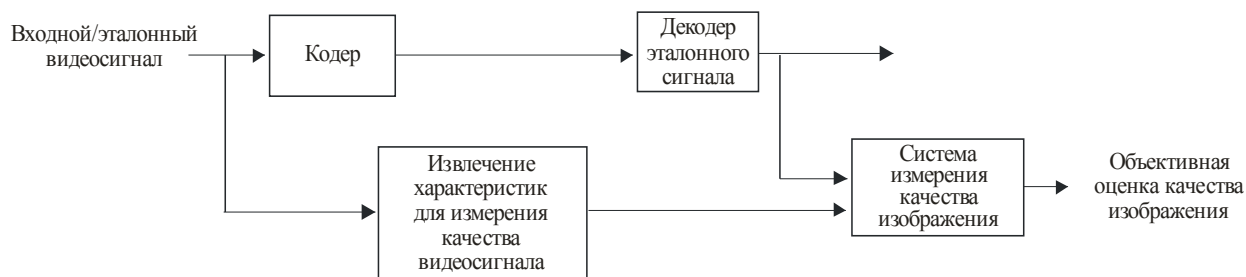
## 6 Описание метода измерения с ухудшенным эталонным сигналом

В случае метода симметричного измерения с ухудшенным эталонным сигналом, предназначенного для целей объективного измерения воспринимаемого качества изображения, оценивается работа систем путем сравнения характеристик неискаженного входного, или эталонного, видеосигнала на входе системы и ухудшенного сигнала на выходе системы (см. рисунок 1).

На рисунке 1 представлен пример применения метода с ухудшенным эталонным сигналом для испытания кодека в лаборатории.

РИСУНОК 1

### Применение метода измерения воспринимаемого качества с ухудшенным эталонным сигналом для испытания кодека в лаборатории



ВТ.1885-01

Сравнение входного и выходного сигналов может потребовать осуществления временного согласования или пространственного согласования, последнее – для компенсации какого-либо вертикального или горизонтального смещения или обрезки изображения. Оно может также потребовать коррекции сдвига или разницы усиления в каналах яркости и цветности. Далее вычисляется объективная оценка качества изображения, как правило, путем применения модели восприятия глазом человека.

Согласование и коррекция усиления называется нормализацией. Этот процесс требуется в силу того, что в большинстве методов с ухудшенным эталонным сигналом осуществляется практически попиксельное сравнение свойств, выделенных из эталонных изображений и обработанных изображений. Примером будет служить расчет пикового отношения сигнал/шум (PSNR). Удаляются только независимые от времени статические изменения в видеосигнале, а динамические изменения, обусловливаемые испытываемой системой, измеряются в рамках расчета оценки качества.

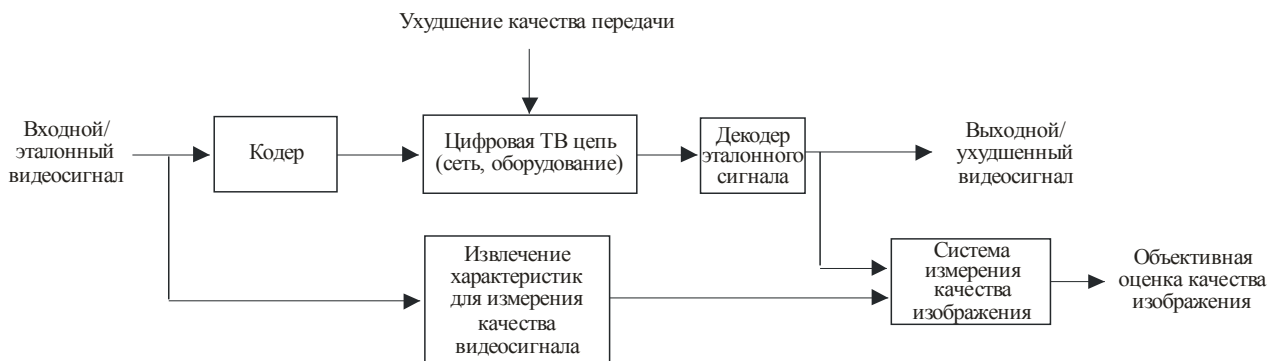
В Рекомендациях МСЭ-T J.244 и МСЭ-T J.144 приведены стандартные методы представления значений, необходимых для нормализации видеосигнала, до осуществления объективной оценки качества. Показатели качества изображения, описанные в Приложении к настоящей Рекомендации, включают соответствующие методы нормализации. В отношении показателей качества изображений, приведенных в Приложении, могут применяться альтернативные методы нормализации при условии, что они обеспечивают требуемую точность нормализации.

В силу того что показатели качества изображения, как правило, базируются на аппроксимациях к зрительным реакциям человека, а не на измерении конкретных артефактов кодирования, они в принципе в равной мере действительны как для аналоговых, так и для цифровых систем. Они также в принципе действительны для цепей, в состав которых входят аналоговые и цифровые системы или каскады систем с цифровым сжатием.

На рисунке 2 представлен пример применения метода с ухудшенным эталонным сигналом для измерения цепи передачи.

РИСУНОК 2

**Применение метода измерения воспринимаемого качества  
с ухудшенным эталонным сигналом для испытания цепи передачи**



ВТ.1885-02

В этом случае декодер эталонного сигнала получает сигналы в разных точках цепи передачи, например декодер может размещаться в какой-либо точке сети, как на рисунке 2, или непосредственно на выходе кодера, как на рисунке 1. Если цифровая сеть передачи является прозрачной, измерение объективной оценки качества изображения в источнике аналогично измерению в любой последующей точке этой цепи.

Общепризнано, что метод с полным эталонным сигналом обеспечивает большую точность измерений воспринимаемого качества изображения. Это метод продемонстрировал возможность высокой корреляции с субъективными оценками, производимыми в соответствии с методами ACR-HR, описанными в Рекомендации МСЭ-T P.910.

## 7 Заключение Группы экспертов по качеству видеосигнала

Исследования измерений воспринимаемого качества изображения проводятся в рамках неофициальной группы, называемой Группой экспертов по качеству видеосигнала (VQEG), которая подотчетна 9-й и 12-й Исследовательским комиссиям МСЭ-T и 6-й Исследовательской комиссии МСЭ-R. В ходе недавно проведенных Группой VQEG испытаний RRNR-TV оценивалась работа предлагаемых алгоритмов измерения воспринимаемого качества изображения с ухудшенным эталонным сигналом для форматов МСЭ-R 601-6.

На основании представленных данных в настоящее время МСЭ-T может рекомендовать шесть методов RR (Модель\_A 15k, Модель\_A 80k, Модель\_A 256k, Модель\_C 80k, Модель\_B 80k (только 525 строк), Модель\_B 256k (только 525 строк)).

Технические описания этих моделей представлены в Приложениях А–С, соответственно. Следует заметить, что порядок следования приложений является абсолютно произвольным и не отражает характеристик прогнозирования качества.



В таблицах 1 и 2 представлены результаты проверок по критерию значимости в рамках испытаний VQEG RRNR-TV. Для формата 525 строк четыре модели (Модель\_А 15к, Модель\_А 80к, Модель\_А 256к, Модель\_С 80к) являются статистически более точными по сравнению с PSNR, и две модели (Модель\_В 80к, Модель\_В 256к) являются статистически эквивалентными PSNR. Следует заметить, что PSNR было рассчитано NTIA с использованием метода поиска полным перебором пределов калибровки. Для формата 625 строк четыре модели (Модель\_А 15к, Модель\_А 80к, Модель\_А 256к, Модель\_С 80к) являются статистически эквивалентными и статистически более точными относительно PSNR.

ТАБЛИЦА 1

**Испытание по критерию значимости для формата 525 строк**

Формат 525 строк	По сравнению с лучшим	По сравнению с PSNR	Корреляция
Модель_А 15к	1	1	0,906
Модель_А 80к	1	1	0,903
Модель_А 256к	1	1	0,903
Модель_С 80к	1	1	0,882
Модель_В 80к	0	1	0,795
Модель_В 256к	0	1	0,803
PSNR_NTIA	0	1	0,826

ПРИМЕЧАНИЕ 1. – "1" в "По сравнению с лучшим" означает, что модель статистически эквивалентна лучшей модели. "0" означает, что модель статистически не эквивалентна лучшей модели. "1" в "По сравнению с PSNR" означает, что модель статистически эквивалентна лучшей модели. "0" означает, что модель статистически не эквивалентна лучшей модели.

ТАБЛИЦА 2

**Испытание по критерию значимости для формата 625 строк**

Формат 625 строк	По сравнению с лучшим	По сравнению с PSNR	Корреляция
Модель_А 15к	1	1	0,894
Модель_А 80к	1	1	0,899
Модель_А 256к	1	1	0,898
Модель_С 80к	1	1	0,866
PSNR_NTIA	0	1	0,857

ПРИМЕЧАНИЕ 1. – "1" в "По сравнению с лучшим" означает, что модель статистически эквивалентна лучшей модели. "0" означает, что модель статистически не эквивалентна лучшей модели. "1" в "По сравнению с PSNR" означает, что модель статистически эквивалентна лучшей модели. "0" означает, что модель статистически не эквивалентна лучшей модели.

В таблицах 3 и 4 представлена справочная информация по характеристикам моделей, использованных в рамках испытаний VQEG RRNR-TV.

ТАБЛИЦА 3

**Справочное описание характеристик моделей, использованных  
в рамках испытаний VQEG RRNR-TV  
(формат 525 строк)**

<b>Формат 525 строк</b>	<b>Корреляция</b>	<b>RMSE</b>	<b>OR</b>
Модель_A 15k	0,906	0,418	0,385
Модель_A 80k	0,903	0,423	0,378
Модель_A 256k	0,903	0,424	0,378
Модель_B 80k	0,795	0,598	0,667
Модель_B 256k	0,803	0,587	0,647
Модель_C 80k	0,882	0,465	0,513
PSNR_NTIA	0,826	0,556	0,571

ТАБЛИЦА 4

**Справочное описание характеристик моделей, использованных  
в рамках испытаний VQEG RRNR-TV  
(формат 625 строк)**

<b>Формат 625 строк</b>	<b>Корреляция</b>	<b>RMSE</b>	<b>OR</b>
Модель_A 15k	0,894	0,524	0,468
Модель_A 80k	0,899	0,513	0,462
Модель_A 256k	0,898	0,516	0,468
Модель_C_80k	0,866	0,585	0,583
PSNR_NTIA	0,857	0,605	0,564

## Приложение А

### **Модель А: Метод с ухудшенным эталонным сигналом, разработанный в университете Йонсей**

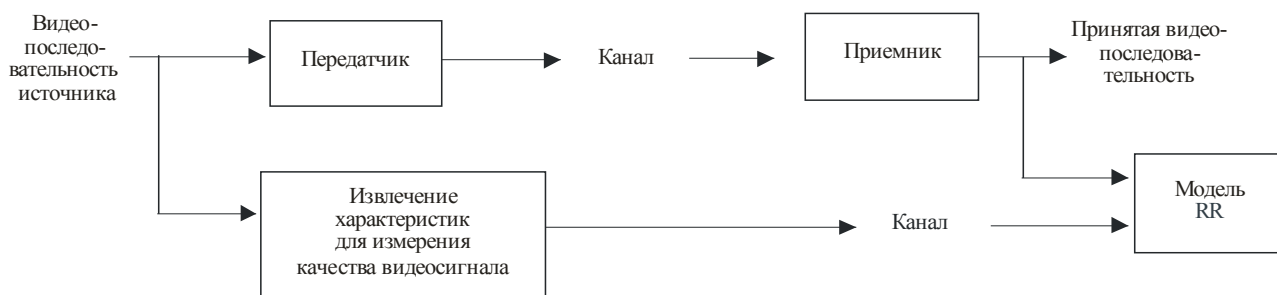
#### **1 Введение**

Несмотря на то что PSNR широко применяется в качестве объективного критерия качества видеосигнала, также отмечается, что этот критерий не обеспечивает точного представления воспринимаемого качества изображения. По результатам анализа восприятия человеком качества изображения было выявлено, что зрение человека чувствительно к ухудшению качества на границах. Другими словами, если пиксели на контурах изображения размыты, эксперты скорее всего дадут низкую оценку изображению, даже при высоком значении PSNR. На основе данного наблюдения были созданы модели с ухудшенным эталонным сигналом, с помощью которых в основном измеряют ухудшение качества на контурах.

На рисунке 3 показано, как работает модель с ухудшенным эталонным сигналом. Свойства, которые будут использованы для измерения качества видеосигнала в точке наблюдения, извлекаются из видеопоследовательности источника и передаются. В таблице 5 приведены значения ширины полосы бокового канала для свойств, которые проходили проверку при испытании VQEG RRNR-TV.

РИСУНОК 3

## Функциональная схема модели с ухудшенным эталонным сигналом



ВТ.1885-03

ТАБЛИЦА 5

## Ширина полосы боковых каналов

Формат изображения	Испытываемая ширина полосы
Формат 525 строк	15 кбит/с, 80 кбит/с, 256 кбит/с
Формат 625 строк	15 кбит/с, 80 кбит/с, 256 кбит/с

## 2 Модели EPSNR с ухудшенным эталонным сигналом

## 2.1 Контурное PSNR

С помощью моделей RR в основном измеряется ухудшение качества на контурах. В этих моделях алгоритм выделения контуров прежде всего применяется к видеопоследовательности источника, с тем чтобы определить контурные пиксели. Затем измеряется ухудшение качества этих контурных пикселей посредством вычисления среднеквадратичной ошибки. Из среднеквадратичной ошибки вычисляется контурное PSNR (EPSNR).

Разрешается использовать любой алгоритм выделения контуров, хотя в результате могут наблюдаться небольшие отличия. Например, для выделения контуров можно использовать любой оператор градиента. Было предложено несколько операторов градиента. Во многих алгоритмах выделения контуров, предусматривающих применение градиентных операторов, сначала вычисляют, применяя градиентные операторы, изображение с горизонтальным градиентом  $g_{\text{горизонтальное}}(m,n)$  и изображение с вертикальным градиентом  $g_{\text{вертикальное}}(m,n)$ . Затем по следующей формуле рассчитывается величина градиента изображения  $g(m,n)$ :

$$g(m,n) = |g_{\text{горизонтальное}}(m,n)| + |g_{\text{вертикальное}}(m,n)|.$$

Наконец, для того чтобы найти контурные пиксели, к величине градиента изображения  $g(m,n)$  применяют операцию определения порога. Говоря иначе, пиксели, значения величины градиентов которых превышают пороговое значение, считаются контурными пикселями.

Рисунки 4–8 служат для иллюстрации этой процедуры. На рисунке 4 показано изображение источника. На рисунке 5 показано изображение с горизонтальным градиентом  $g_{\text{горизонтальное}}(m,n)$ , которое получено в результате применения к исходному изображению на рисунке 4 оператора горизонтального градиента. На рисунке 6 показано изображение с вертикальным градиентом  $g_{\text{вертикальное}}(m,n)$ , которое получено в результате применения к изображению источника на рисунке 4 оператора вертикального градиента. На рисунке 7 показано изображение градиента величины (контурное изображение), а на рисунке 8 показано контурное изображение в двоичном коде (изображение маски), полученное в результате применения к изображению градиента величины на рисунке 7 определения порога.

РИСУНОК 4

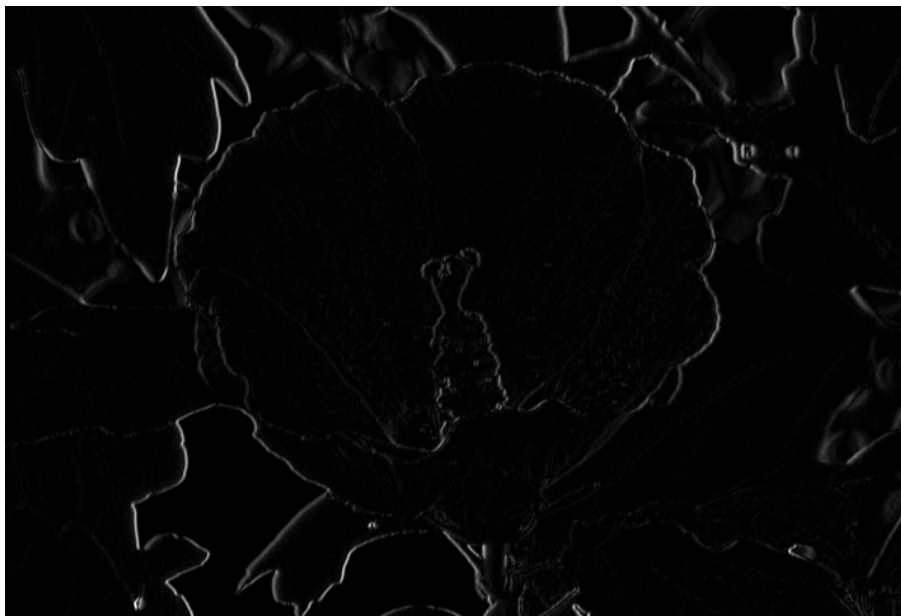
Изображение источника (исходное изображение)



ВТ.1885-04

РИСУНОК 5

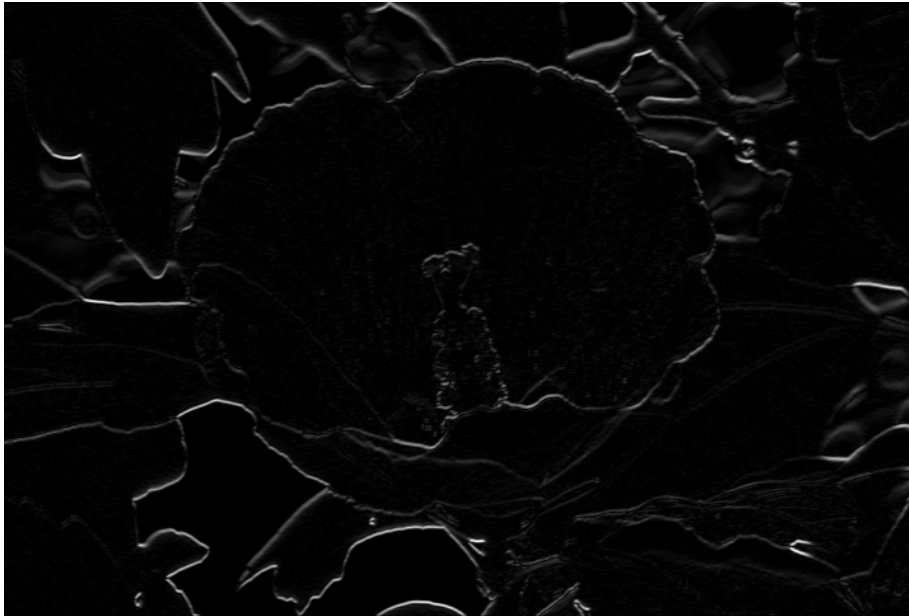
Изображение с горизонтальным градиентом, которое получено в результате применения к изображению источника на рисунке 4 оператора горизонтального градиента



ВТ.1885-05

РИСУНОК 6

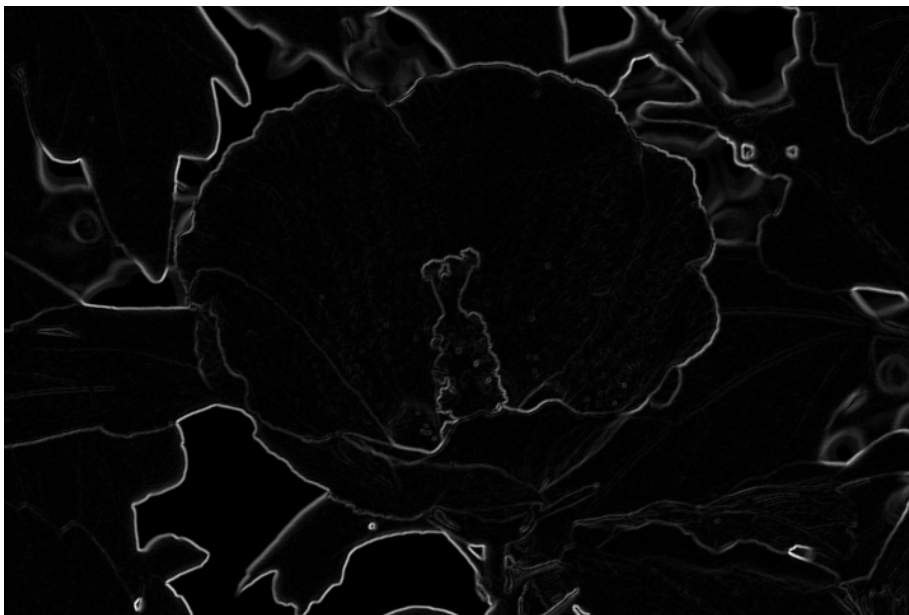
Изображение с вертикальным градиентом, которое получено в результате применения к изображению источника на рисунке 4 оператора вертикального градиента



ВТ.1885-06

РИСУНОК 7

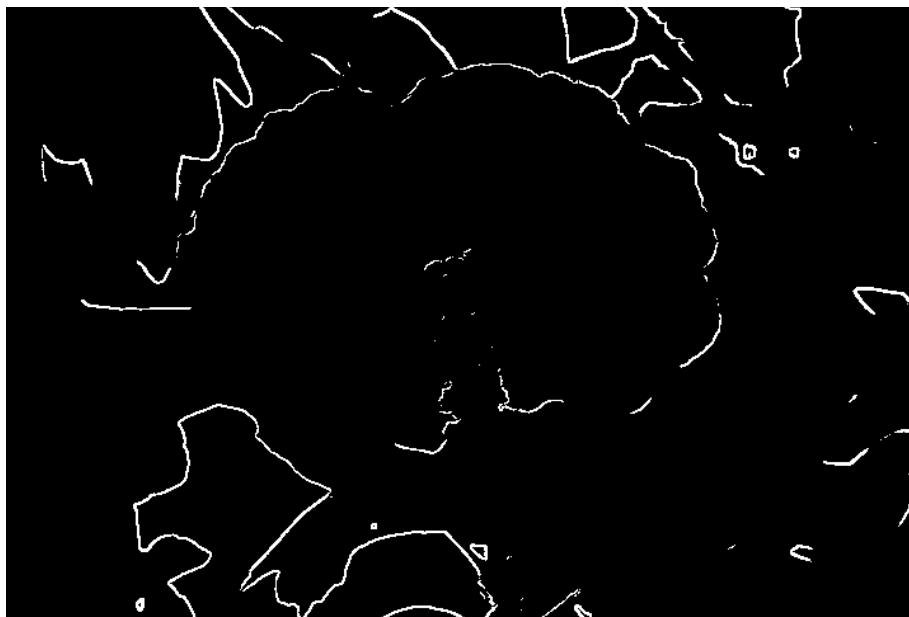
Изображение градиента величины



ВТ.1885-07

РИСУНОК 8

Контурное изображение в двоичном формате (изображение маски), полученное в результате применения к изображению градиента величины на рисунке 7 определения порога



ВТ.1885-08

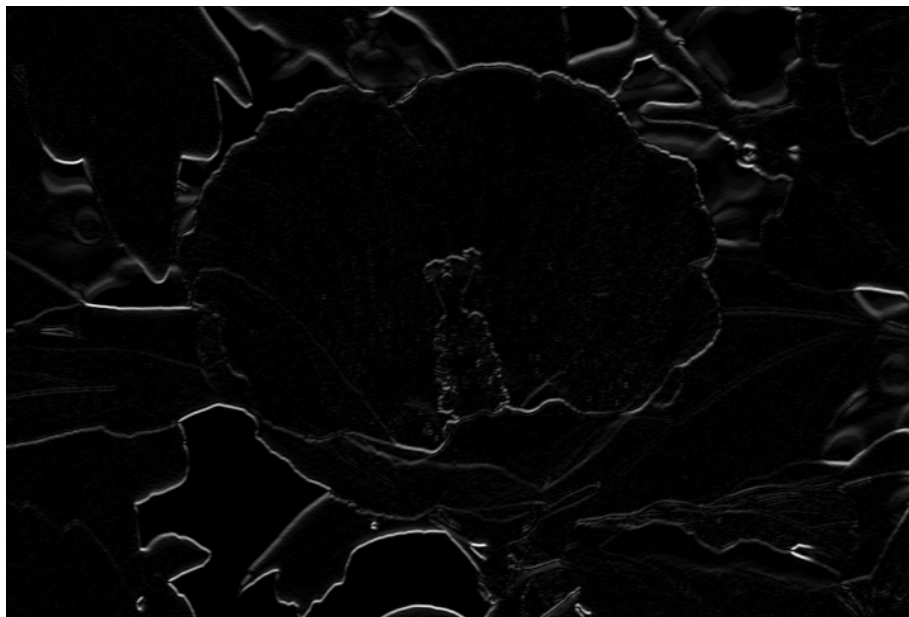
С другой стороны, для определения контуров можно использовать измененную процедуру. Например, сначала к изображению источника можно применить оператор вертикального градиента, создав изображение с вертикальным градиентом. Затем к изображению с вертикальным градиентом применяется оператор горизонтального градиента, создавая измененное изображение с последовательным градиентом (изображение с горизонтальным и вертикальным градиентом). Наконец, чтобы найти контурные пиксели, к измененному изображению с последовательным градиентом можно применить операцию определения порога. Иными словами, пиксели измененного изображения с последовательным градиентом, градиенты которых превышают пороговое значение, считаются контурными пикселями. Рисунки 9–12 служат для иллюстрации измененной процедуры. На рисунке 9 показано изображение с вертикальным градиентом  $g_{\text{вертикальное}}(m,n)$ , которое получено в результате применения к исходному изображению на рисунке 4 оператора вертикального градиента. На рисунке 10 показано измененное изображение с последовательным градиентом (изображение с горизонтальным и вертикальным градиентом), которое получено в результате применения к изображению с вертикальным градиентом на рисунке 9 оператора горизонтального градиента. На рисунке 11 показано контурное изображение в двоичном коде (изображение маски), полученное в результате применения к измененному изображению с последовательным градиентом на рисунке 10 определения порога.

Следует заметить, что оба метода могут рассматриваться как алгоритмы определения контуров. Можно выбрать любой алгоритм определения контуров в зависимости от свойств изображений и алгоритмов сжатия. Однако некоторые методы могут быть эффективнее других.

Таким образом, в этой модели сначала применяется оператор определения контуров, создавая контурные изображения (рисунки 7–10). Затем создается изображение маски (контурное изображение в двоичном коде) посредством применения к контурному изображению определения порога (рисунки 8–11). То есть пиксели контурного изображения, значения которых меньше порогового  $t_e$ , устанавливаются в нулевое значение, а пиксели с равным или большим значением порога устанавливаются в ненулевое значение. На рисунках 8 и 11 показаны изображения маски. Так как видеоизображения могут просматриваться как последовательность кадров или полей, описанная выше процедура может применяться к каждому кадру или полю видеоизображения. Так как эта модель может использоваться для видеоизображения на основе полей или на основе кадров, для указания на поле или кадр будет использоваться термин "изображение".

РИСУНОК 9

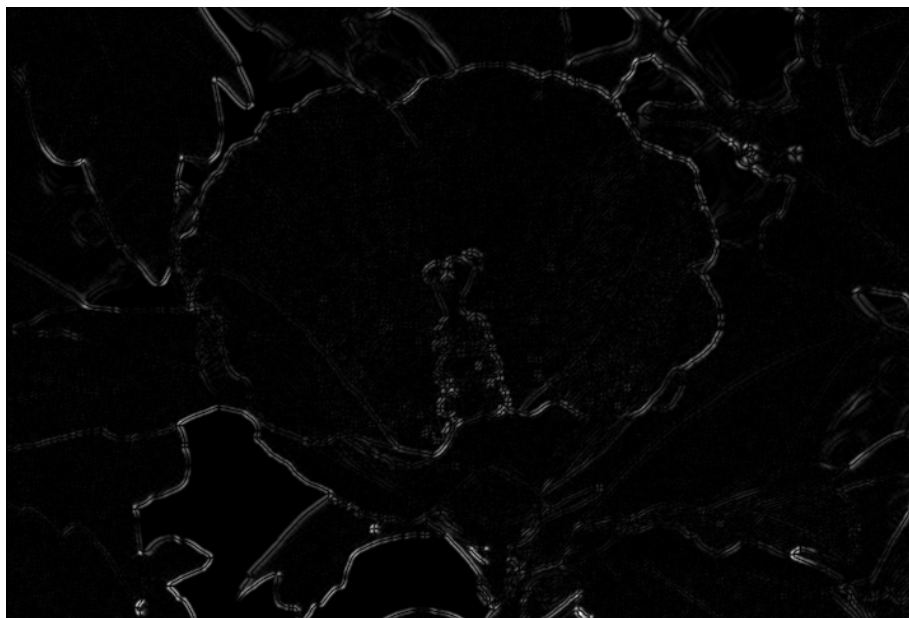
Изображение с вертикальным градиентом, которое получено в результате применения к изображению источника на рисунке 4 оператора вертикального градиента



ВТ.1885-09

РИСУНОК 10

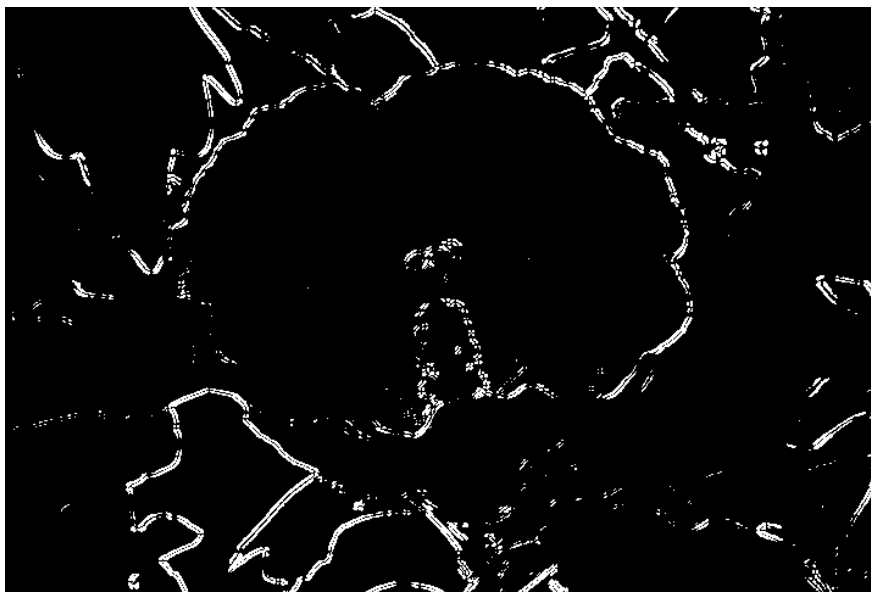
Измененное изображение с последовательным градиентом (изображение с горизонтальным и вертикальным градиентом), которое получено в результате применения к изображению с вертикальным градиентом на рисунке 9 оператора горизонтального градиента



ВТ.1885-10

РИСУНОК 11

Контурное изображение в двоичном коде (изображение маски), полученное в результате применения к измененному изображению с последовательным градиентом на рисунке 10 определения порога



ВТ.1885-11

## 2.2 Выбор свойств из видеопоследовательностей источника

Поскольку данная модель является моделью RR, из каждого изображения видеопоследовательности источника необходимо извлечь свойства. В модели EPSNR RR из каждого изображения извлекается определенное количество контурных пикселей. Затем местоположение и значения пикселя кодируются и пересылаются. Однако для некоторых видеопоследовательностей при использовании фиксированного порогового значения количество контурных пикселей может быть весьма небольшим. В сценарии худшего случая оно может равняться нулю (пустые изображения или изображения с очень низкой частотой). Для решения этой проблемы, в случае если количество контурных пикселей изображения меньше заданного значения, пользователь может уменьшать пороговое значение до тех пор, пока количество контурных пикселей не превысит заданного значения. С другой стороны, можно выбрать контурные пиксели, которые соответствуют наибольшим значениям изображений с горизонтальным и вертикальным градиентом. Если в кадре нет контурных пикселей, например пустые изображения, можно случайным образом выбрать необходимое количество пикселей или пропустить этот кадр. Например, если из каждого кадра следует выбрать 10 контурных пикселей, можно отсортировать пиксели изображений с горизонтальным и вертикальным градиентом в соответствии с их значениями и выбрать 10 наибольших значений. Однако в результате данной процедуры в одинаковых местоположениях может оказаться большое число контурных пикселей. Для решения этой проблемы сначала можно выбрать несколько раз необходимое количество пикселей изображения с горизонтальным или вертикальным градиентом, а затем в случайном порядке выбрать из отобранных пикселей изображения с горизонтальным или вертикальным градиентом необходимое количество контурных пикселей. В проверенных в ходе испытаний VQEG RRNR-TV моделях необходимое количество контурных пикселей выбрано случайным образом из большого множества контурных пикселей. Это множество контурных пикселей получено путем применения к градиентному изображению операции определения порога.

В моделях RR EPSNR местоположение и значения контурных пикселей кодируются после применения фильтра Гаусса нижних частот к выбранным местоположениям пикселей. Несмотря на то что в рамках испытаний VQEG RRNR-TV использовался фильтр Гаусса нижних частот ( $5 \times 3$ ), в зависимости от формата изображения могут использоваться разные фильтры нижних частот. Следует заметить, что в процессе кодирования возможно применение обрезки. Во избежание выбора контурных пикселей в обрезанных областях, модель выбирает контурные пиксели в центральной области (см. рисунок 12). В таблице 6 приведены размеры после обрезки. Также в таблице 6 приведено количество битов, необходимых для кодирования местоположения и значения контурного пикселя.



ТАБЛИЦА 6

## Требования по битам для каждого контурного пикселя

Формат изображения	Размер	Размер после обрезки	Битов для местоположения	Битов для значения пикселя	Всего битов на пиксель
525	720 × 486	656 × 438	19	8	27
625	720 × 576	656 × 528	19	8	27

РИСУНОК 12

## Пример обрезки и центральной области



ВТ.1885-12

Модель выбирает контурные пиксели из каждого кадра в соответствии с разрешенной шириной полосы (см. таблицу 5). В таблице 7 приведено количество контурных пикселей на кадр, которые могут быть переданы для испытываемой ширины полосы.

ТАБЛИЦА 7

## Количество контурных пикселей на кадр

Формат изображения	15 кбит/с	80 кбит/с	256 кбит/с
525	16	74	238
625	20	92	286

РИСУНОК 13  
Блок-схема модели



ВТ.1885-13

### 2.3 Пространственная/временная регистрация и коррекция усиления/сдвига

До вычисления разницы между контурными пикселями в видеопоследовательности источника и пикселями в обработанной видеопоследовательности, которая является видеопоследовательностью, принятой приемником, модель вначале применяет пространственную/временную регистрацию и коррекцию усиления/сдвига. Использовался метод калибровки (Приложение В), представленный в Рекомендации МСЭ-Т J.244. Для передачи свойств усиления и сдвига согласно Рекомендации МСЭ-Т J.244 (Приложение В) в ходе испытаний VQEG RRNR-TV использовались 30% доступной ширины полосы. Поскольку видеопоследовательность является чересстрочной, метод калибровки применяется три раза: четные поля, нечетные поля и смешанные кадры. Если разница между ошибкой четного поля (PSNR) и ошибкой нечетного поля превышала пороговое значение, использовались результаты регистрации (x-сдвиг, y-сдвиг) с меньшими PSNR. В противном случае использовались результаты регистрации со смешанными кадрами. В ходе испытаний VQEG RRNR-TV пороговое значение было установлено равным 2 дБ.

В точке наблюдения обработанную видеопоследовательность следует согласовать с контурными пикселями, извлеченными из видеопоследовательности источника. Однако, если ширина полосы боковых каналов невелика, доступно лишь небольшое количество контурных пикселей видеопоследовательности источника (см. рисунок 14). Следовательно, временная регистрация может быть неточна, если эта временная регистрация осуществляется с использованием единственного кадра (см. рисунок 15). Для решения данной проблемы в модели используется окно для временной регистрации. Вместо использования отдельного кадра обработанной видеопоследовательности, модель, для отыскания оптимального временного сдвига, создает окно, которое состоит из нескольких соседних кадров. Эта процедура показана на рисунке 16. Среднеквадратичная ошибка в пределах окна вычисляется следующим образом:

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

где:

$MSE_{window}$ : среднеквадратичная ошибка окна;

$E_{SRC}(i)$ : контурный пиксель в пределах окна, который имеет соответствующий пиксель в обработанной видеопоследовательности;

$E_{PVS}(i)$ : пиксель в обработанной видеопоследовательности, соответствующий контурному пикселю;

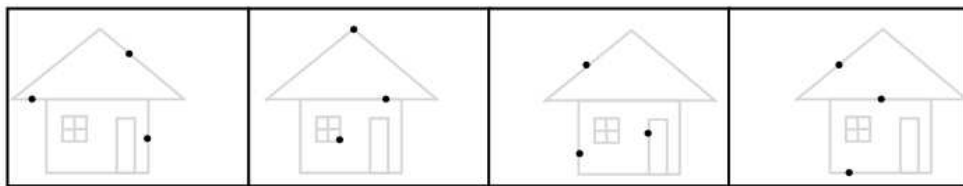
$N_{win}$ : общее количество контурных пикселей, использованных для вычисления  $MSE_{window}$ .

Эта среднеквадратичная ошибка окна используется в качестве разницы между кадром обработанной видеопоследовательности и соответствующим кадром видеопоследовательности источника.

Размер окна может быть определен по результатам анализа характера обработанной видеопоследовательности. Для стандартного применения рекомендуется окно, соответствующее двум секундам. В качестве альтернативы могут применяться окна с разными размерами, и использоваться будет наилучшее, которое обеспечивает наименьшее значение среднеквадратичной ошибки. Кроме того, для учета пропуска кадров в результате ошибок передачи (см. рисунок 20) могут использоваться разные центры окна.

РИСУНОК 14

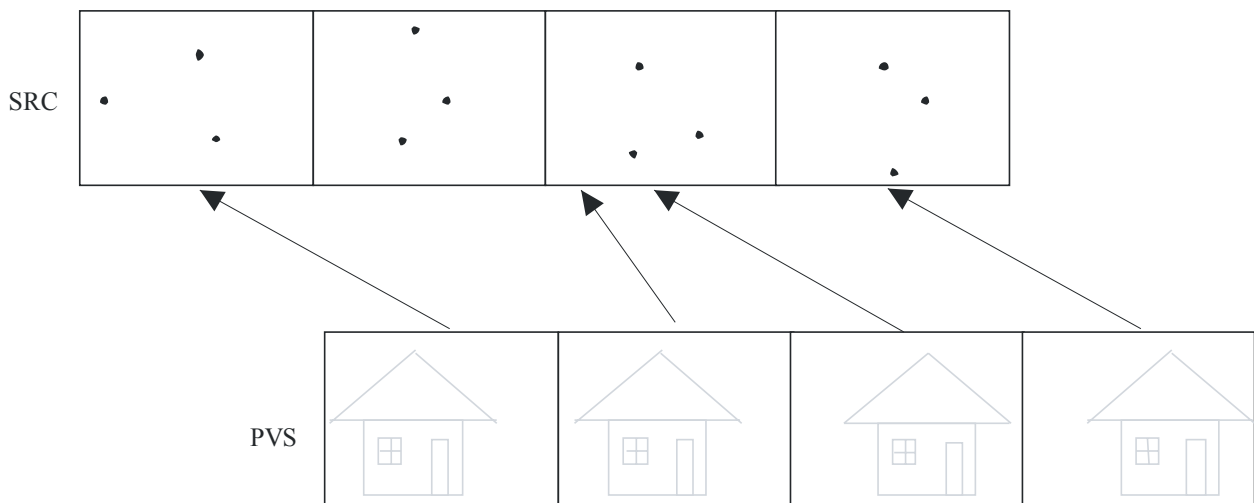
Выбор контурных пикселей в видеопоследовательности источника



ВТ.1885-14

РИСУНОК 15

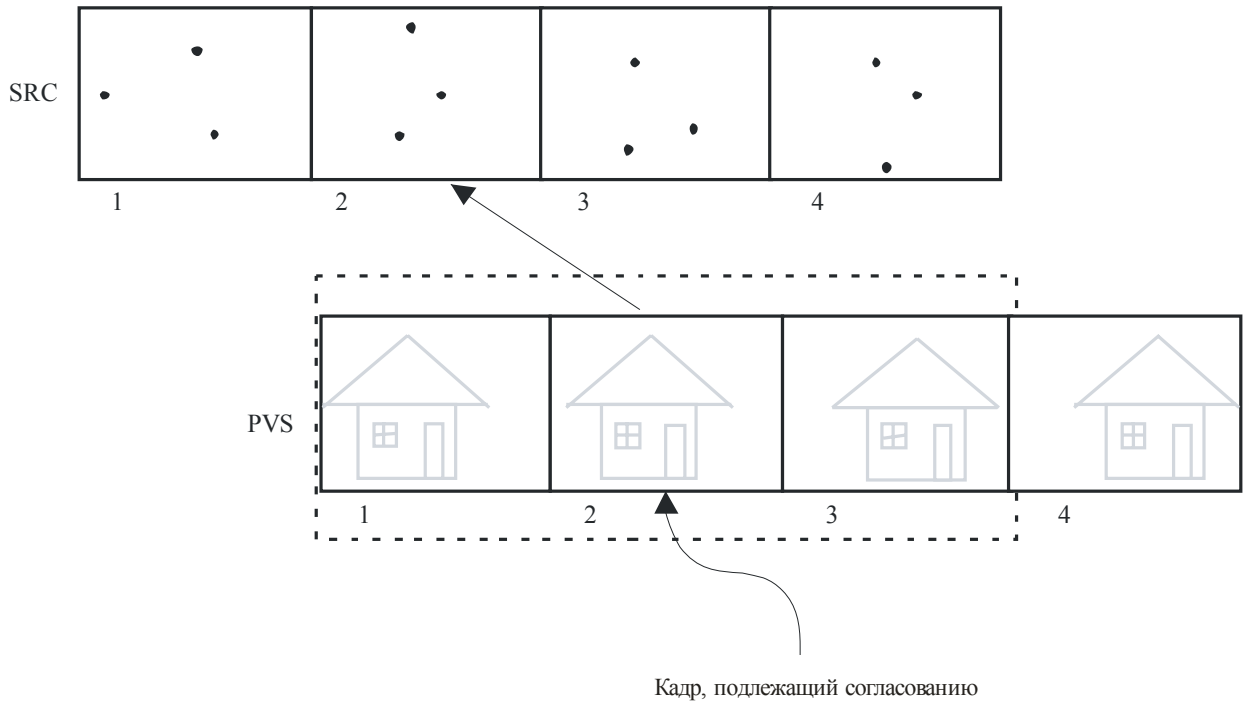
Согласование обработанной видеопоследовательности с контурными пикселями видеопоследовательности источника



ВТ.1885-15

РИСУНОК 16

Согласование обработанной видеопоследовательности с контурными пикселями с использованием окна

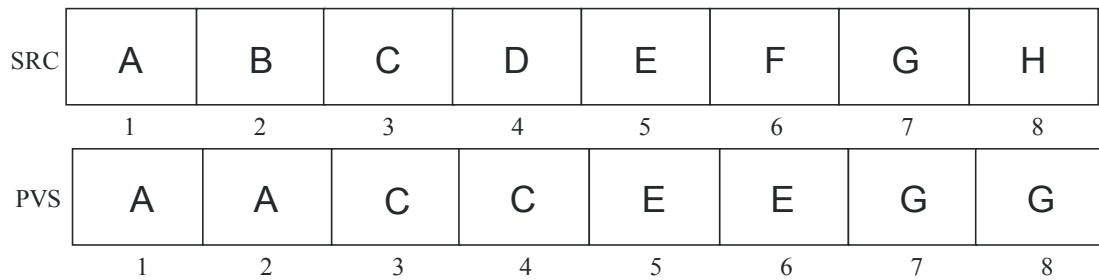


ВТ.1885-16

Если видеопоследовательность источника кодирована с высоким коэффициентом сжатия, кодер может уменьшать количество кадров в секунду, а обработанная видеопоследовательность содержит повторяющиеся кадры (см. рисунок 17). На рисунке 17 обработанная видеопоследовательность не содержит кадров, соответствующих некоторым кадрам видеопоследовательности источника (2-й, 4-й, 6-й, 8-й кадры). В этом случае для расчета среднеквадратичной ошибки модель не использует повторяющиеся кадры. Иначе говоря, эта модель выполняет временную регистрацию, используя первый кадр (действительный кадр) каждого повторяющегося блока. Поэтому на рисунке 18 для временной регистрации используются только три кадра (3-й, 5-й, 7-й кадры) в пределах окна.

РИСУНОК 17

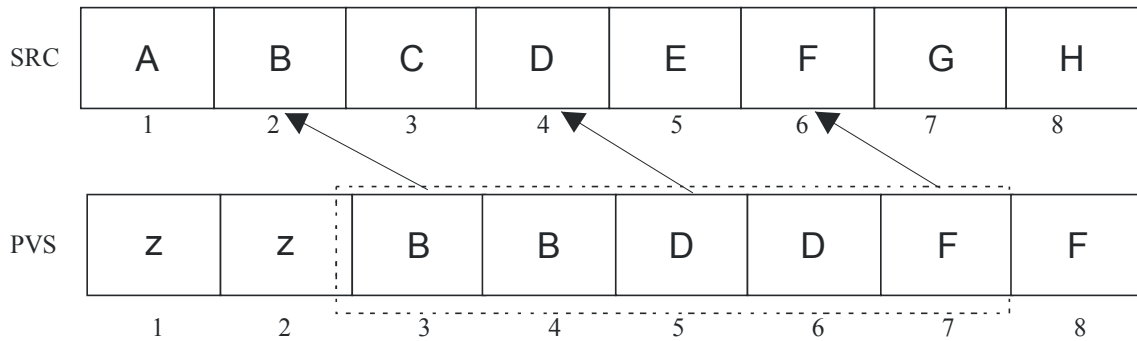
Пример повторяющихся кадров



ВТ.1885-17

РИСУНОК 18

## Обработка повторяющихся кадров



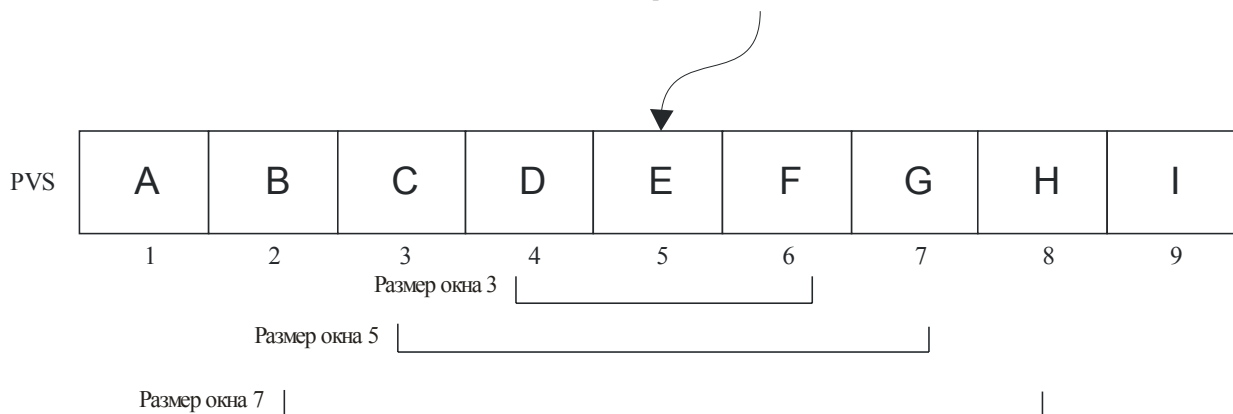
ВТ.1885-18

Вероятно получение обработанной видеопоследовательности с нерегулярным повторением кадров, что может привести к тому, что метод временной регистрации с использованием окна даст неточные результаты. Для решения этой проблемы после временной регистрации с использованием окна возможно локально корректировать каждый кадр окна в пределах заданного значения (например,  $\pm 1$ ), как показано на рисунке 21. Затем для расчета EPSNR используется локальная коррекция, которая обеспечивает минимальную MSE.

РИСУНОК 19

## Окна разных размеров

Кадр, подлежащий согласованию

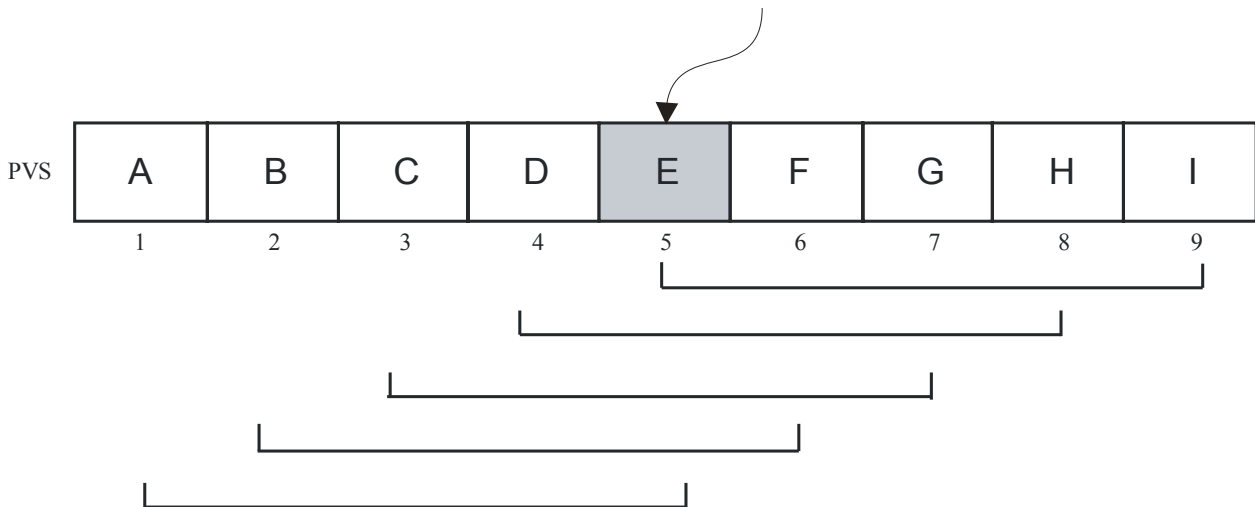


ВТ.1885-19

РИСУНОК 20

## Центры окон

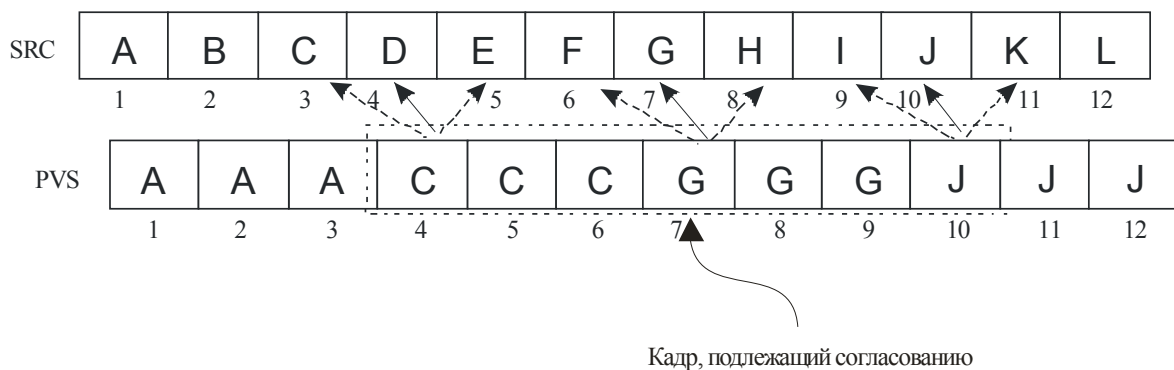
Кадр, подлежащий согласованию



ВТ.1885-20

РИСУНОК 21

## Локальная коррекция временной регистрации с использованием окна



ВТ.1885-21

## 2.4 Вычисление EPSNR и пост-обработка

После выполнения временной регистрации, вычисляется среднее значение разницы между контурными пикселями видеопоследовательности источника и соответствующими пикселями обработанной видеопоследовательности, которое может рассматриваться в качестве контурной среднеквадратичной ошибки обработанной видеопоследовательности ( $MSE_{edge}$ ). В завершение по следующей формуле вычисляется EPSNR (контурное PSNR):

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

где  $P$  – пиковое значение изображения.

## 1 Фиксация кадров

При кодировании мультимедийного видеосигнала может наблюдаться повторение кадров из-за уменьшения частоты кадров и фиксация кадров из-за ошибок передачи, которые ухудшают воспринимаемое качество изображения. В целях устранения этого эффекта до вычисления EPSNR в модели применяется следующая коррекция:

$$MSE_{frozen\_frame\_considered} = MSE_{edge} \times \frac{K \times N_{total\_frame}}{N_{total\_frame} - N_{total\_frozen\_frame}},$$

где:

$MSE_{frozen\_frame\_considered}$ : среднеквадратичная ошибка, которая учитывает повторяющиеся и фиксированные кадры;

$N_{total\_frame}$ : общее количество кадров,  $N_{total\_frozen\_frame}$ ;

$K$ : постоянная.

В проверяемой в рамках испытаний VQEG RRNR-TV модели величина  $K$  была установлена равной 1.

## 2 Высокая частота и движение с высокой скоростью

Если видеопоследовательность содержит большой объем высокой частоты и движений с высокой скоростью, воспринимаемое качество при той же MSE обычно повышается. Для учета этого эффекта определяются нормализованная межкадровая разница (NFD) и нормализованная энергия высокой частоты (NHFE) по следующей формуле:

$$NFD = \frac{FD}{\text{среднее значение энергии на пиксель}},$$

где  $FD = \frac{1}{N_F} \sum_i \sum_{k=1}^{height} \sum_{j=1}^{width} (Frame_i[j, k] - Frame_{i-1}[j, k])^2$  и  $N_F$  – количество кадров, использованное

для суммирования. Следует заметить, что при расчете FD три наибольших значения межкадровой разницы не учитываются, с тем чтобы исключить из расчетов среднего значения межкадровой разницы смену сцены, принимая длительность видеопоследовательности, равную 8 с. Нормализованная энергия высокой частоты (NHFE) вычисляется путем расчета средних значений энергии высокой частоты (см. рисунок 22) после применения преобразования Фурье 2D:

$$NHFE = \frac{\text{средние значения энергии высокой частоты}}{\text{среднее значение энергии на пиксель}},$$

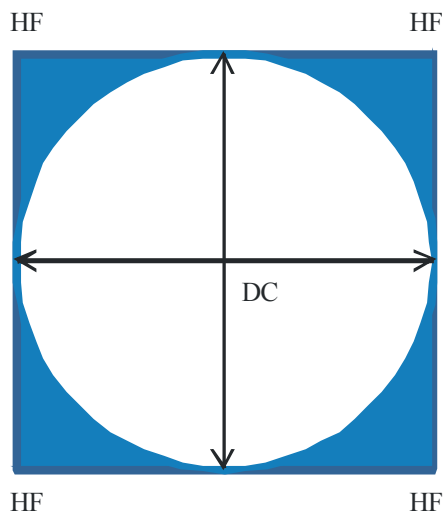
В завершение используются следующие уравнения:

```
IF(SNFD > 0.35 && SNHFE > 2.5) {
    IF(EPSNR < 20) EPSNR = EPSNR+3
    ELSE IF(EPSNR < 35) EPSNR = EPSNR+5
}
ELSE IF((SNFD > 0.2 && SNHFE > 1.5) || (SNFD>0.27) && SNHFE > 1.3)) {
    IF(28 < EPSNR < 40) EPSNR = EPSNR + 3
    IF(EPSNR > 40) EPSNR = 40
}
```

где SNFD – NFD источника, а SNHFE – NHFE источника. Следует заметить, что SNFD и SNHFE рассчитываются по SRC и передаются как часть данных, характеризующих свойства (1 байт на каждое).

РИСУНОК 22

**Вычисление нормализованной высокочастотной энергии.**  
 Значения энергии высокой частоты вычисляются по затененной области



ВТ.1885-22

### 3 Размытость

Для учета явлений размывания используются следующие уравнения:

```

IF (NHFE/SNHFE < 0.5)
    IF(EPSNR>26)    EPSNR = 26
ELSE IF (NHFE/SNHFE < 0.6)
    IF(EPSNR>32)    EPSNR = 32
ELSE IF (NHFE/SNHFE < 0.7)
    IF(EPSNR>36)    EPSNR = 36
ELSE IF (NHFE/SNHFE > 1.2)
    IF(EPSNR>23)    EPSNR = 23
ELSE IF (NHFE/SNHFE > 1.1)
    IF (EPSNR>25)    EPSNR = 25 ,
  
```

где NHFE – PVS NHFE.

### 4 Блочность

Для учета эффектов блочности рассчитывается среднее значение разницы по столбцам. При условии вычислений по модулю 8 оценка блочности для  $i$ -го кадра рассчитывается следующим образом:

$$Blk[i] = \frac{\text{наибольшая разность по столбцам}}{\text{следующая после наибольшей разности по столбцам}}$$

Окончательная оценка блочности (*Блочность*) рассчитывается путем усреднения оценок блочности кадров:

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

В завершение применяются следующие уравнения:

```

IF(BLOCKING > 1.4) {
    IF (20≤EPSNR<25) EPSNR = EPSNR-1.086094*BLOCKING-0.601316
    ELSE IF (EPSNR<30) EPSNR = EPSNR-0.577891*BLOCKING-3.158586
    ELSE IF (EPSNR<35) EPSNR = EPSNR-0.223573*BLOCKING-3.125441
}
  
```



### 5 Максимальная длительность фиксированных кадров

Ошибки передачи могут обуславливать длительные фиксированные кадры. Для учета длительных фиксированных кадров используются следующие уравнения:

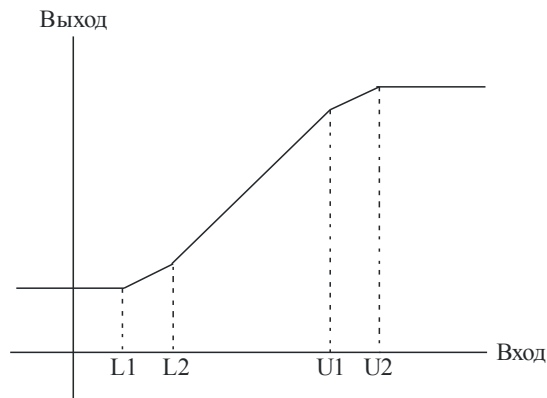
$$\begin{aligned} & \text{IF}(\text{MAX\_FREEZE} > 22 \text{ AND } \text{EPSNR} > 28) \text{ EPSNR} = 28 \\ & \text{ELSE IF}(\text{MAX\_FREEZE} > 10 \text{ AND } \text{EPSNR} > 34) \text{ EPSNR} = 34, \end{aligned}$$

где MAX\_FREEZE – наибольшая длительность фиксированных кадров. Необходимо заметить, что если длительность видеопоследовательности не равна 8 с, следует использовать другие пороговые значения.

### 6 Подгонка с помощью кусочно-линейной функции

Если EPSNR превышает определенное значение, происходит насыщение воспринимаемого качества. В этом случае возможно установить верхнюю границу EPSNR. Кроме того, если желательно линейное соотношение EPSNR и DMOS (средняя субъективная оценка разницы), возможно применить кусочно-линейную функцию, как показано на рисунке 23. В модели, проверяемой в рамках испытаний VQEG RRNR-TV, верхняя граница устанавливалась равной 48, а нижняя граница – равной 15.

РИСУНОК 23  
Кусочно-линейная функция для линейного соотношения  
EPSNR и DMOS



ВТ.1885-23

Модели EPSNR с ухудшенным эталонным сигналом, применяемые для объективного измерения качества изображения, базируются на ухудшении по контурам. Эти модели могут быть реализованы в реальном времени при умеренном использовании вычислительных ресурсов. Такие модели хорошо подходят для применений, требующих контроля качества изображения в реальном времени при условии доступности боковых каналов.

## Приложение В

### Модель В: метод с ухудшенным эталонным сигналом, разработанный фирмой NEC

В настоящем Приложении представлено полное функциональное описание модели RR. В модели RR на сторону клиента вместо значений пикселей для отдельных блоков пикселей заданного размера передаются значения активности. Качество изображения оценивается на основе разницы активности между эталонным каналом источника (SRC) и обработанной видеопоследовательностью (PVS). Точность оценки должны повысить психовизуальные весовые коэффициенты.

Данная модель не требует пространственной регистрации и регистрации усиления и сдвига, требующей вычислительных ресурсов. Кроме того, такая модель может быть реализована с помощью программы, состоящей из 30 строк, и программы, состоящей из 250 строк, установленных на сторонах сервера и клиента соответственно. Следовательно, она пригодна для контроля качества изображения в реальном времени в радиовещательных службах, где основными преимуществами являются невысокий уровень сложности, а также простота реализации.

#### 1 Резюме

В модели RR на сторону клиента вместо значений пикселей для отдельных блоков пикселей заданного размера передаются значения активности. Качество изображения оценивается на основе разницы активности между SRC и PVS. Точность оценки должны повысить психовизуальные весовые коэффициенты.

Данная модель не требует пространственной регистрации и регистрации усиления и сдвига, требующей вычислительных ресурсов. Кроме того, такая модель может быть реализована с помощью программы, состоящей из 30 строк, и программы, состоящей из 250 строк, установленных на сторонах сервера и клиента соответственно. Следовательно, она пригодна для контроля качества изображения в реальном времени в радиовещательных службах, где основными преимуществами являются невысокий уровень сложности, а также простота реализации.

#### 2 Определения

**Активность (Activity)** – среднее значение абсолютной разницы между каждым значением яркости и средним значением яркости для блока заданного размера.

**Блок (Block)** – матрица пикселей  $M \times N$  ( $M$ -столбцов на  $N$ -рядов).

**Кадр (Frame)** – одно полное телевизионное изображение.

**Усиление (Gain)** – множительный масштабный коэффициент, применяемый гипотетическим эталонным трактом (HRC) ко всем пикселям отдельной плоскости изображения (например, яркость, цветность). Усиление сигнала яркости обычно называют контрастностью.

**Гипотетическая эталонная цепь (Hypothetical Reference Circuit) (HRC)** – испытываемая видеосистема, например кодек или цифровая система передачи видеоизображений.

**Яркость (Luminance) (Y)** – часть видеосигнала, который в основном переносит информацию яркости (т. е. монохромная часть изображения).

**Национальный комитет по телевизионным системам (National Television Systems Committee) (NTSC)** – система аналогового цветного телевидения с полным сигналом с 525-строчным кадром [1].

**Сдвиг или сдвиг уровня (Offset or level offset)** – аддитивный фактор, применяемый гипотетическим эталонным трактом (HRC) ко всем пикселям отдельной плоскости изображения (например, яркость, цветность). Сдвиг сигнала яркости обычно называют яркостью.

**Пиковое отношение сигнал/шум (Peak signal-to-noise ratio) (PSNR)** – отношение максимальной возможной мощности сигнала и мощности искажающего сигнал шума.

**Фазопеременная линия (Phase-Altering Line) (PAL)** – система аналогового цветного телевидения с полным сигналом с 625-строчным кадром.

**Растровая развертка (Raster scan)** – отображение прямоугольного двухкоординатного поля в однокоординатное поле таким образом, что первая точка однокоординатного поля берется из первого верхнего ряда двухкоординатного поля, сканируемого слева направо, далее аналогично из второго, третьего и т. д. рядов поля (перемещаясь вниз), каждый из которых сканируется слева направо.

**Эталонный сигнал в уменьшенной полосе частот (ухудшенный эталонный сигнал) (Reduced-reference) (RR)** – методика измерения качества изображения, в которой используются узкополосные свойства, извлеченные из исходных или обработанных видеопотоков, в противоположность использованию эталонного видеосигнала в полной полосе частот, которое требует полной информации об исходных и обработанных видеопотоках [2]. Методики с использованием эталонного сигнала в уменьшенной полосе характеризуются преимуществами для контроля сквозного качества в процессе эксплуатации, поскольку информация об эталонном сигнале в уменьшенной полосе легко передается по повсеместно распространенным сетям электросвязи.

**Изучаемая область (Region of interest) (ROI)** – решетка изображения (определенная в прямоугольных координатах), которая используется для определения конкретной подобласти поля или кадра видеоизображения.

**Сцена (Scene)** – последовательность видеок кадров.

**Пространственная регистрация (Spatial registration)** – процесс, используемый для оценки и коррекции пространственных сдвигов обработанной видеопоследовательности относительно исходной видеопоследовательности.

**Временная регистрация (Temporal registration)** – процесс, используемый для оценки и коррекции временных сдвигов (например, задержки изображения) обработанной видеопоследовательности относительно исходной видеопоследовательности.

**Метрика качества изображения (Video Quality Metric) (VQM)** – общая мера ухудшения качества изображения. VQM представляется в виде одного числа и имеет номинальный диапазон от нуля до единицы, где нуль означает отсутствие воспринимаемого ухудшения, а единица – максимальное воспринимаемое ухудшение.

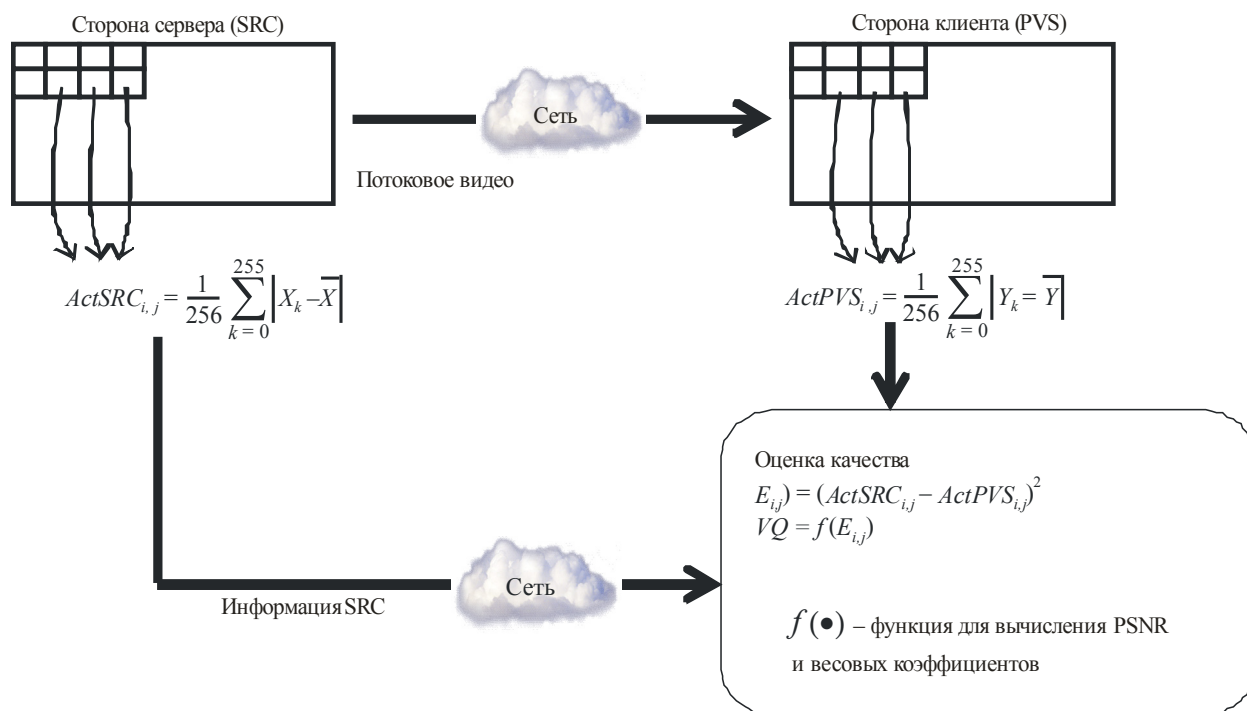
### 3 Общее описание вычисления метрики качества изображения

Модель RR передает на сторону клиента значения активности отдельных блоков пикселей заданных размеров. Эти значения указывают дисперсию значений яркости в данном блоке. На рисунке 1 в общем виде представлена модель RR. Как показано на рисунке 24, качество изображения оценивается на основе разницы активности между SRC и PVS. Кроме того, для повышения точности оценки к разнице активности применяются психовизуальные весовые коэффициенты. Оценка качества изображения осуществляется в следующей последовательности:

- 1 На стороне сервера рассчитывается значение активности для каждого блока пикселей яркости  $16 \times 16$  SRC. Затем все значения активности передаются на сторону клиента. Значение активности отдельного блока определяется как средняя абсолютная разница значения этого блока и среднего значения.
- 2 На стороне клиента рассчитываются соответствующие значения активности относительно PVS.
- 3 На стороне клиента каждый блок первоначально оценивается по его квадратичной ошибке, т. е. квадратичной разнице между значениями активности SRC и PVS.
- 4 Психовизуальные весовые коэффициенты применяются к среднеквадратичным ошибкам в блоках с высокими пространственными компонентами частоты, конкретным цветом, большим значением межкадровой разницы и сменой сцены.
- 5 Предварительная оценка качества изображения выводится по сумме взвешенных квадратичных ошибок способом, аналогичным расчету PSNR.
- 6 Эта оценка модифицируется для отражения фатальных ухудшений вследствие блочности и локальных искажений. В результате модифицированная оценка отражает измеренное качество изображения PVS в модели RR.

РИСУНОК 24

## Оценка качества изображения на основе разницы активности



ВТ.1885-24

## 4 Детальное описание алгоритма

## 4.1 Сторона сервера

- 1 В каждом кадре через одну секунду после верхней точки видеопоследовательности пиксели яркости SRC разделяются на блоки пикселей  $16 \times 16$ . В течение первой секунды информация SRC не передается, так как для зрения человека сложно определить ухудшение качества изображения в сценах сразу же с первого кадра.
- 2 Для каждого блока, кроме блоков по периметру кадра, рассчитываются значения активности (активность SRC:  $ActSRC_{i,j}$ ). На рисунке 25 показаны блоки, значения активности которых рассчитываются и передаются. Активность SRC рассчитывается следующим образом:

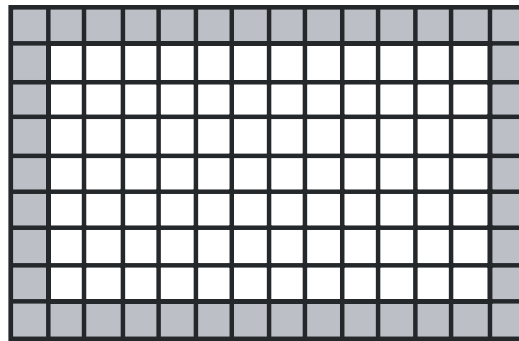
$$ActSRC_{i,j} = \frac{1}{256} \sum_{k=0}^{255} |X_k - \bar{X}|,$$

где  $X_k$  – значение яркости в блоке заданного размера SRC,  $\bar{X}$  – его среднее значение,  $i$  – номер кадра,  $j$  – номер блока в кадре.

- 3 Значения активности, которые выражаются восемью битами на блок, передаются на сторону клиента в порядке растровой развертки через одну секунду с верхней точки видеопоследовательности. При передаче информации SRC со скоростью 256 кбит/с значения активности рассчитываются во всех кадрах. Если скорость снижается до 80 кбит/с, значения активности рассчитываются в каждом четырех кадрах.

РИСУНОК 25

## Блоки с передачей и без передачи значений активности



Блоки с передачей значений активности (16 × 16 пикселей)



Блоки без передачи значений активности (16 × 16 пикселей)

ВТ.1885-25

## 4.2 Сторона клиента

### 4.2.1 Расчет среднеквадратичной ошибки значений активности

- 1 В каждом кадре через одну секунду с верхней точки видеопоследовательности пиксели яркости PVS разделяются на блоки пикселей 16 × 16.
- 2 Для каждого блока, кроме блоков по периметру кадра, рассчитываются значения активности (активность PVS:  $ActPVS_{i,j}$ ). При передаче информации SRC со скоростью 256 кбит/с значения активности рассчитываются во всех кадрах. Если скорость передачи информации SRC снижается до 80 кбит/с, значения активности рассчитываются в каждом четырех кадрах.

$$ActPVS_{i,j} = \frac{1}{256} \sum_{k=0}^{255} |Y_k - \bar{Y}|,$$

где:

 $Y_k$ : значение яркости в блоке заданного размера PVS; $\bar{Y}$ : его среднее значение; $i$ : номер кадра; $j$ : номер блока в кадре.

- 3 Квадратичные значения расхождения активности SRC и активности PVS рассчитываются следующим образом:

$$E_{i,j} = (ActSRC_{i,j} - ActPVS_{i,j})^2.$$

### 4.2.2 Психовизуальное взвешивание для квадратичной ошибки

Для учета характеристик зрения человека к  $E_{i,j}$  применяется три типа взвешивания: взвешивание разницы пространственной частоты, взвешивание разницы в конкретной цветовой области и взвешивание межкадровой разницы яркости.

### 1 Взвешивание разницы пространственной частоты

Для этого типа взвешивания используются весовой коэффициент  $W_{SF}$  и пороговое значение  $Th_{SF}$ . (значения  $W_{SF}$  и  $Th_{SF}$  см. в таблице 8.)

$$E_{ij} \Leftarrow \begin{cases} E_{ij} \times W_{SF}, & ActPVS_{i,j} > Th_{SF} \\ E_{ij}, & иначе \end{cases}.$$

### 2 Взвешивание разницы в конкретной цветовой области

Для любого данного блока и прилегающих к нему восьми блоков, если число пикселей ( $NumROI Pixels$ ) находится в диапазоне  $48 \leq Y \leq 224$ ,  $104 \leq Cb \leq 125$ , а  $135 \leq Cr \leq 171$  превышает пороговое значение, выполняется следующее взвешивание с использованием весового коэффициента  $W_{CR}$  и порогового значения  $Th_{CR}$ :

$$E_{ij} \Leftarrow \begin{cases} E_{ij} \times W_{CR}, & NumROI Pixels > Th_{CR} \\ E_{ij}, & иначе \end{cases}.$$

Значения  $W_{CR}$  и  $Th_{CR}$  см. в таблице 8.

### 3 Взвешивание межкадровой разницы яркости

Рассчитывается средняя абсолютная разница ( $MAD_{i,j}$ ) яркости данного блока и предшествующего кадра:

$$MAD_{i,j} = \frac{1}{256} \sum_{k=0}^{255} |Y_k - Y'_k|,$$

где  $Y_k$  – значение яркости в блоке пикселей  $16 \times 16$  PVS и  $Y'_k$  – значение яркости в той же позиции в предшествующем кадре.

Выполняется следующая операция взвешивания с использованием весовых коэффициентов  $W_{MAD1}$ ,  $W_{MAD2}$  и пороговых значений  $Th_{MAD1}$ ,  $Th_{MAD2}$ .

$$E_{ij} \Leftarrow \begin{cases} E_{ij} \times W_{MAD1}, & MAD_{i,j} > Th_{MAD1} \\ E_{ij} \times W_{MAD2}, & MAD_{i,j} \leq Th_{MAD2} \\ E_{ij}, & иначе \end{cases}.$$

Значения  $W_{MAD1}$ ,  $W_{MAD2}$ ,  $Th_{MAD1}$  и  $Th_{MAD2}$  см. в таблице 8.

#### 4.2.3 Взвешивание в случае определения смены сцены

Для каждого кадра рассчитывается среднее  $MAD_{i,j}$  ( $MADAve_i$ ):

$$MADAve_i = \frac{1}{M} \sum_{j=0}^{M-1} MAD_{i,j},$$

где  $M$  – количество блоков в кадре.

Если значение  $MAD_{Ave_i}$  превышает пороговое значение  $Th_{SC}$ , считается, что произошла смена сцены. Если определяется смена сцена, значение  $E_{ij}$  устанавливается равным 0 для 15 кадров, следующих после смены сцены (SceneChange).

$$SceneChange = \begin{cases} TRUE, & MAD_{Ave_i} > Th_{SC} \\ FALSE, & \text{иначе} \end{cases}.$$

$$E_{ij} \Leftarrow \begin{cases} E_{ij} \times W_{SC} & 15 \text{ кадров после } SceneChange = TRUE \\ E_{ij}, & \text{иначе} \end{cases}.$$

Значения  $W_{SC}$  и  $Th_{SC}$  см. в таблице 8.

#### 4.2.4 PSNR на основе квадратичной ошибки активности

PSNR рассчитывается на основании разницы активности:

$$VQ = 10 \times \log_{10} \frac{255 \times 255}{E_{Ave}}$$

$$E_{Ave} = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} E_{i,j},$$

где  $N$  и  $M$  – количество кадров и блоков, использованных для расчета PSNR.

#### 4.2.5 Взвешивание артефактов блочности

Для этой операции взвешивания используются весовой коэффициент  $W_{BL}$ , пороговое значение  $Th_{BL}$  и информация об уровне блочности  $BL_{Ave}$ . (Значения  $W_{BL}$  и  $Th_{BL}$  см. в таблице 8.)

$$VQ \Leftarrow \begin{cases} VQ \times W_{BL}, & BL_{Ave} > Th_{BL} \\ VQ, & \text{иначе} \end{cases}.$$

$BL_{Ave}$  рассчитывается в следующей последовательности:

*Шаг 1:* Рассчитываются значения активности для блоков пикселей  $8 \times 8$  в PVS. Как показано на рисунке 26, среднее значение ( $Act_{Ave}$ ) двух значений активности в горизонтально согласованных блоках ( $ActBlock_1, ActBlock_2$ ) рассчитывается следующим образом:

$$Act_{Ave} = \frac{1}{2} (ActBlock_1 + ActBlock_2).$$

*Шаг 2:* Рассчитывается абсолютная разница значений яркости на границе между двумя блоками. Как показано на рисунке 26,  $Y_{1,0}$  и  $Y_{2,0}$  представляют значения яркости в левом и правом блоках вдоль границы. Среднее значение абсолютной разницы яркости,  $DiffBound$ , выражается следующим образом:

$$DiffBound = \frac{1}{8} \sum_{i=0}^7 |Y_{1,i} - Y_{2,i}|.$$

*Шаг 3:* Определяется уровень блочности ( $BL_{i,j}$ ) по отношению  $DiffBound$  и  $Act_{Ave}$ , т. е.:

$$BL_{i,j} = \frac{DiffBound}{Act_{Ave} + 1}.$$

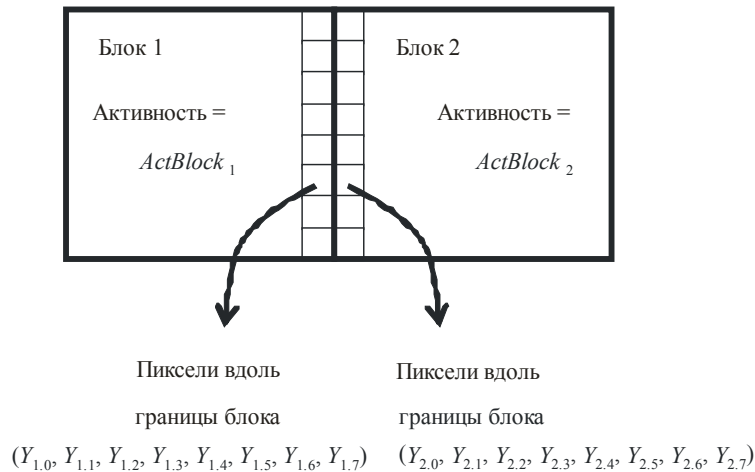
Шаг 4: Среднее значение  $BL$  рассчитывается следующим образом:

$$BL_{Ave} = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} BL_{i,j}.$$

Для крайних правых блоков в кадрах значение  $BL_{i,j}$  устанавливается равным нулю. Если  $BL_{Ave}$  превышает заранее установленное пороговое значение, считается, что данная видеопоследовательность имеет высокий уровень блочности, и для расчета значения качества изображения применяется взвешивание.

РИСУНОК 26

Значения пикселей и активности для вычисления уровня блочности



ВТ.1885-26

#### 4.2.6 Взвешивание локальных артефактов искажения

Для этой операции взвешивания используются весовой коэффициент  $W_{LI}$ , пороговое значение  $Th_{LI}$  и локальное искажение  $LI$ . (Значения  $W_{LI}$  и  $Th_{LI}$  см. в таблице 8.)

$$VQ \Leftarrow \begin{cases} VQ \times W_{LI}, & LI > Th_{LI} \\ VQ, & \text{иначе} \end{cases}.$$

$LI$  рассчитывается в следующей последовательности. Для обнаружения локальных искажений вследствие ошибок передачи используется разница дисперсии активности.

- 1 Для данного блока и восьми прилегающих к нему блоков рассчитывается дисперсия активности как для SRC ( $ActVar_{SRC}$ ), так и для PVS ( $ActVar_{PVS}$ ) и вычисляется абсолютная разница этих значений дисперсии:

$$\Delta ActVar = |ActVar_{SRC} - ActVar_{PVS}|.$$

- 2 Для каждого кадра рассчитывается среднее этих значений абсолютной разницы.
- 3  $LI$  рассчитывается как отношение максимума ( $\Delta ActVar_{Max}$ ) к минимуму ( $\Delta ActVar_{Min}$ ) среднего значения.

$$LI = \begin{cases} \Delta ActVar_{Min} / \Delta ActVar_{Max} & \Delta ActVar_{Max} \neq 0 \\ 1 & \Delta ActVar_{Max} = 0. \end{cases}$$

$VQ$  представляет оценку качества изображения.



#### 4.2.7 Параметры для операций взвешивания

В таблице 8 приведены значения параметров для операций взвешивания. Эти значения определяются в рамках предварительного эксперимента с квалификационным набором данных.

ТАБЛИЦА 8  
Параметры взвешивания

Тип операции взвешивания	Значения параметров	
Взвешивание пространственной частоты	$W_{SF}$	0,36
	$Th_{SF}$	25
Взвешивание определенного цвета	$W_{CR}$	4,0
	$Th_{CR}$	175
Взвешивание межкадровой разницы	$W_{MAD1}$	0,06
	$Th_{MAD1}$	17
	$W_{MAD2}$	25
	$Th_{MAD2}$	13
Определение смены сцены	$W_{SC}$	0,0
	$Th_{SC}$	35
Взвешивание блочности	$W_{BL}$	0,870
	$Th_{BL}$	1,0
Взвешивание локальных искажений	$W_{LI}$	0,870
	$Th_{LI}$	1,67

#### 4.2.8 Регистрация

##### 1 Пространственная регистрация

Для модели RR не требуется какая-либо пространственная регистрация. Это обусловливается тем, что квадратичная ошибка рассчитывается по значениям активности, которые более устойчивы к пространственным сдвигам, чем значения, рассчитываемые по значениям пикселей.

##### 2 Регистрация усиления и сдвига

Для модели RR не требуется какая-либо регистрация усиления и сдвига. Значения активности по своему существу не подвержены сдвигу (т. е. компоненты DC) и нечувствительны к усилению.

##### 3 Временная регистрация

Последовательность PVS разделяется на односекундные подпоследовательности. Для каждой подпоследовательности рассчитываются среднеквадратичные ошибки активности при пяти вариантах задержки до  $\pm 2$  кадров SRC. В результате в качестве среднеквадратичной ошибки в каждой подпоследовательности используется минимальное значение среднеквадратичных ошибок. Задержка, результатом которой является эта минимальная среднеквадратичная ошибка, корректируется как временная регистрация.

#### 5 Пример программы

Ниже приведен пример программы на языке C для модели RR.

### 5.1 Общая программа для стороны сервера и стороны клиента

// Расчет значения активности

```
unsigned int CalcActivitybyRect(unsigned char * lpBuffer, int nWidth, int iRectWidth, int iRectHeight)
{
    // lpBuffer: Кадровый буфер яркости
    // nWidth: Ширина кадрового буфера
    // iRectWidth: Ширина прямоугольника для расчета значения активности.
    // iHeightWidth: Высота прямоугольника для расчета значения активности.
    unsigned int i, j, nTmp, nSum;
    unsigned char *pSrc;

    pSrc = lpBuffer; nSum = 0;
    for (j = 0; j < iRectHeight; j++){
        for (i = 0; i < iRectWidth; i++){
            nSum += pSrc[i];
        }
        pSrc += nWidth;
    }
    nSum /= (iRectWidth*iRectHeight);

    pSrc = lpBuffer; nTmp = 0;
    for (j = 0; j < iRectHeight; j++){
        for (i = 0; i < iRectWidth; i++){
            nTmp += abs(pSrc[i] - nSum);
        }
        pSrc += nWidth;
    }
    return nTmp/iRectWidth/iRectHeight;
}
```

### 5.2 Сторона сервера

// Сторона сервера

int nStart = 30; // номер кадра для начала передачи (30 или 25)

int nMaxFrame = 240; // количество всех кадров изображения (240 или 200)

int nFrameIncrement = 1; // 1 для 256 кбит/с, 4 для 80 кбит/с

void ReadOneFrame(unsigned char, int, unsigned char \*, int, int); // функция для считывания данных одного кадра

int nRim = 16 // 16 или 32 (использовать 32 во избежание проблем в HRC9)

// nWidth: Ширина кадрового буфера

// nHeight: Высота кадрового буфера

// lpSrc: Кадровый буфер

```
for(int nFrame = nStart; nFrame < nMaxFrame; nFrame+=nFrameIncrement){
    ReadOneFrame(SRC_file_name, nFrame, lpSrc, nWidth, nHeight);
    for (j = 16; j<nHeight-32; j+=16) {
        for (i = nRim; i < nWidth- nRim; i+=16) {
            lpOrg = lpSrc + i + j * nWidth;
            nActSrc = CalcActivitybyRect(lpOrg, nWidth, 16, 16);
            // OutputSRCInfo(nActSrc); // Вывод или передача информации SRC
        }
    }
}
```

### 5.3 Сторона клиента

// Сторона клиента

int nStart = 30; // номер кадра для начала передачи (30 или 25)

int nMaxFrame = 240; // количество всех кадров изображения (240 или 200)

int nFrameIncrement = 1; // 1 для 256 кбит/с, 4 для 80 кбит/с

int nFrameRate = 30; //30 или 25

void ReadOneFrame(unsigned char, int, unsigned char \*\*, int, int); // функция для считывания данных одного кадра

void ReadRRData(unsigned char, int, unsigned char \*); // функция для считывания данных RR

// nWidth: Ширина кадрового буфера

// nHeight: Высота кадрового буфера

// lpPvsByte[3]: Кадровый буфер (0:Y, 1:Cb, 2:Cr)

// lpRRData: Буфер данных RR

// double ddActivityDifference[][]: Сохранение разницы активности

```

// double ddActivityVariance[]: Сохранение дисперсии активности
// double ddBlock[]: Сохранение уровня блочности
// int nSceneChange: Определение смены сцены

for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){ // Изменение временного согласования
    for(int nFrame = 0; nFrame < nMaxFrames; nFrame++){
        if(nFrame+nTemporalAlign >= nMaxFrames || nFrame+nTemporalAlign < 0){
            continue;
        }
        ReadOneFrame(PVS_file_name, nFrame+nTemporalAlign, lpPvsByte, nWidth, nHeight);
        if(((nFrame-(nFrameRate+nStart)) % nFrameIncrement) == 0
            && nFrame >= nStart ){
            ReadRRData(RR_file_name, nFrame, lpRRData);
            ddActivityDifference[nTemporalAlign+2][nFrame]
                = RRCalcObjectiveScore(lpPvsByte, lpRRData, nWidth, nHeight);
            ddActivityVariance[nTemporalAlign+2][nFrame] = gnActVar;
        }else{
            ddActivityDifference[nTemporalAlign+2][nFrame] = 0.0;
            ddActivityVariance[nTemporalAlign+2][nFrame] = 0.0;
        }
        // Уровень блочности
        if(nTemporalAlign ==0){
            ddBlock[nFrame] = BlockinessLevelEstimation(lpPvsByte[0], nWidth, nHeight);
        }
        // Копия пикселей для расчета межкадровой разницы
        memcpy(lpPrev, lpPvsByte[0], sizeof(char)*nWidth*nHeight);
        if(nSceneChange){
            nSceneChange--;
        }
    }
}

double ddSum[8][5]; // Сумма разницы активности для каждой секунды
double ddActVarSum[8][5]; // Сумма дисперсии активности для каждой секунды
double ddActVarMax[8][5]; // Максимум суммы дисперсии активности
double ddActVarMin[8][5]; // Минимум суммы дисперсии активности
int nnMin[8];
int nnNumFrames[8][5];
#define LARGENUMBER 100000
for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){
    for(int j=0;j<8;j++){ // для каждой одной секунды
        nnNumFrames[j][nTemporalAlign+2] = 0;
        ddActVarMax[j][nTemporalAlign+2] = 0.0;
        ddActVarMin[j][nTemporalAlign+2] = LARGENUMBER;
        ddActVarSum[j][nTemporalAlign+2] = 0.0;
        ddSum[j][nTemporalAlign+2] = 0.0;
        for(int i=nFrameRate*j;i< (j+1)*nFrameRate; i++){
            if(ddActivityDifference[nTemporalAlign+2][i]){
                ddSum[j][nTemporalAlign+2] += ddActivityDifference[nTemporalAlign+2][i];
                nnNumFrames[j][nTemporalAlign+2]++;
            }
            ddActVarSum[j][nTemporalAlign+2] += ddActivityVariance[nTemporalAlign+2][i];
            if(ddActivityVariance[nTemporalAlign+2][i]){
                if(ddActivityVariance[nTemporalAlign+2][i] >
                    ddActVarMax[j][nTemporalAlign+2]){
                    ddActVarMax[j][nTemporalAlign+2] =
                        ddActivityVariance[nTemporalAlign+2][i];
                }
                if(ddActivityVariance[nTemporalAlign+2][i] <
                    ddActVarMin[j][nTemporalAlign+2]){
                    ddActVarMin[j][nTemporalAlign+2] =
                        ddActivityVariance[nTemporalAlign+2][i];
                }
            }
        }
    }
}
}

// Расчет уровня локальных искажений

```

```

double dSum = 0.0;
double dActMax = 0.0;
double dActMin = LARGENUMBER;
int nNumFrames = 0;
for(int j=1; j<8; j++){
    double dMin = LARGENUMBER;
    double dMinSum = LARGENUMBER;
    for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){
        if(ddActVarSum[j][nTemporalAlign+2] < dMin){
            dMin = ddActVarSum[j][nTemporalAlign+2];
            dMinSum = ddSum[j][nTemporalAlign+2];
            nnMin[j] = nTemporalAlign+2;
        }
    }
    dSum += dMinSum;
    nNumFrames += nnNumFrames[j][nnMin[j]];
    if(ddActVarMax[j][nnMin[j]] > dActMax){
        dActMax = ddActVarMax[j][nnMin[j]];
    }
    if(ddActVarMin[j][nnMin[j]] < dActMin){
        dActMin = ddActVarMin[j][nnMin[j]];
    }
}
double dTransError = dActMax/dActMin;

// Расчет уровня блочности
double dBlockinessLevel = 0.0;
for(int i=0; i<nMaxFrames; i++){
    dBlockinessLevel += ddBlock[i];
}
dBlockinessLevel = dSumBlock / (double)(nMaxFrames-nFrameRate);

// Расчет оценки качества изображения
if(nNumFrames && nNumberOfBlocks && dSum){
    dSum = dSum / (double)(nNumFrames)/(double)nNumberOfBlocks;
    dSum = 10*log10(255.0*255.0/dSum); //PSNR на основе разницы активности
    if(dBlockinessLevel > dBlokinessTh){
        dSum /= nBlockinessWeighting; // Взвешивание уровня блочности
    }
    if(dTransError > nErrorTh){
        dSum /= nErrorWeightin; // Взвешивание ошибок передачи
    }
}
return dSum;
}

-----
// Расчет значения MAD
unsigned int CalcMAD(unsigned char *lpSrc, unsigned char *lpSrc2, int nWidth, int nHeight)
{
    // lpSrc: Кадровый буфер текущего кадра
    // lpSrc2: Кадровый буфер предшествующего кадра
    unsigned int nSum = 0;
    for (y = 0; y < nHeight; y++) {
        for (x = 0; x < nWidth; x++) {
            nSrc = lpSrc[x + y*nWidth];
            nSrc2 = lpSrc2[x + y*nWidth];
            nSum += abs(nSrc - nSrc2);
        }
    }
    return nSum/nWidth/nHeight;
}

// Расчет среднеквадратичной ошибки с взвешиванием
double RRCalcObjectiveScore(unsigned char *lpBuffer[], unsigned char *lpRRData, int nWidth, int nHeight)
{
    int i, j, nActSrc, nActSrc2, nY, nCb, nCr, nYMin, nYMax, nCbMin, nCbMax, nCrMin, nCrMax;
    int nMBX, nMBY, nMB, nStart;
    unsigned int nMAD;
    double e2, dMADFrame;
    unsigned char *lpRec, *lpRecCb, *lpRecCr, *lpPrev1;

```

```

int nRim = 16 // 16 или 32 (использовать 32 во избежание проблем в HRC9)

nYMin = 48; nYMax = 224; nCbMin = 104; nCbMax = 125; nCrMin = 135; nCrMax = 171;
nMB = nMBY = nStart = 0;
e2 = dMADFrame = 0.0;

for (j=16+nStart; j<ilmageHeight-32; j+=16) {
    nMBX = 0;
    for (i= nRim; i<nWidth- nRim; i+=16) {
        lpRec = lpBuffer[0] + i + j * nWidth;
        lpRecCb = lpBuffer[1] + i/2 + (j/2) * nWidth/2;
        lpRecCr = lpBuffer[2] + i/2 + (j/2) * nWidth/2;
        lpPrev1 = lpPrev + i + j * nWidth;

        nActSrc = lpRRData[nMB]; // Активность SRC
        nActSrc2 = CalcActivitybyRect(lpRec, nWidth, 0, 16, 16); // Активность PVS
        nActArray[nMB] = (double)nActSrc;
        nActArray2[nMB] = (double)nActSrc2;
        e2 += (double)(nActSrc - nActSrc2)*(nActSrc - nActSrc2); // Среднеквадратичная ошибка

        nMAD = CalcMAD(lpRec, lpPrev1, 16, 16, nWidth); // Межкадровая разница
        dMADFrame += (double)nMAD;

        int nNumROIPIxels=0;
        for(int jj=-16;jj<32; jj++){
            for(int ii=-16;ii<32; ii++){
                nY = lpRec[ii];
                nCb = lpRecCb[ii/2];
                nCr = lpRecCr[ii/2];
                if(nY >= nYMin && nY <= nYMax
                    && nCb >= nCbMin && nCb <= nCbMax
                    && nCr >= nCrMin && nCr <= nCrMax){
                    nNumROIPIxels++;
                }
            }
            lpRec += nWidth;
            if((jj & 1) == 1){
                lpRecCb += nWidth/2;
                lpRecCr += nWidth/2;
            }
        }

        // Взвешивание пространственной частоты
        if(nActSrc2 > gdwActThHigh){
            e2 *= dActWeightingHigh;
        }else if(nActSrc2 > gdwActThLow){
            e2 *= dActWeightingMiddle;
        }else {
            e2 *= dActWeightingLow;
        }

        // Взвешивание конкретной цветовой области
        if(nNumROIPIxels > dwROITh){
            e2 *= dROIWeighting;
        }

        // Взвешивание межкадровой разницы
        if(nMAD > dwMADThHigh){
            e2 *= dMADWeightingHigh;
        }else if(nMAD > dwMADThLow){
            e2 *= dMADWeightingMiddle;
        }else {
            e2 *= dMADWeightingLow;
        }
        nMB++;      nMBX++;
    }
}
nMBY++;
}

```

```

// Расчет дисперсии активности по девяти прилегающим блокам.
double nSumActSrc, nSumActPvs, nActVar, nActVar2;
nSumActSrc = nSumActPvs = nActVar = nActVar2 = 0.0;
gnActVar = 0.0;
for (j=1; j<nMBY-1; j++) {
    for (i=1; i<nMBX-1; i++) {
        nSumActSrc = 0.0;
        nSumActPvs = 0.0;
        for(int jj=-1; jj<2; jj++){
            for(int ii=-1; ii<2; ii++){
                nSumActSrc += nActArray[i+ii+nMBX*(j+jj)];
                nSumActPvs += nActArray2[i+ii+nMBX*(j+jj)];
            }
        }
        nSumActSrc /= 9.0;
        nSumActPvs /= 9.0;

        nActVar = 0.0;
        nActVar2 = 0.0;
        for(int jj=-1; jj<2; jj++){
            for(int ii=-1; ii<2; ii++){
                nActVar += (nActArray[i+ii+nMBX*(j+jj)]-nSumActSrc)*(nActArray[i+ii+nMBX*(j+jj)]-nSumActSrc);
                nActVar2 += (nActArray2[i+ii+nMBX*(j+jj)]-
                    nSumActPvs)*(nActArray2[i+ii+nMBX*(j+jj)]-
                    nSumActPvs);
            }
        }
        nActVar /= 9.0;
        nActVar2 /= 9.0;
        gnActVar += abs(nActVar- nActVar2);
    }
}
// Средняя дисперсия активности кадра
gnActVar = gnActVar/(double)(nMBY-2)/(double)(nMBY-2);

// Определение смены сцены
if(dMADFrame/ nMB > 35){
    nSceneChange = 15;
}
if(nSceneChange){
    e2 = 0.0;
}
gnFrame++;

return e2;
}

```

// Расчет уровня блочности

double BlockinessLevelEstimation(unsigned char \*lpBuffer, int nWidth, int nHeight)

```

{
    // lpBuffer: кадровый буфер
    int i, j, nActSrc, nActSrc2, nDiff, nMB;
    unsigned char *lpRec = lpBuffer;
    double dBlock=0.0;

    nMB = 0;
    for (j=0; j<nHeight-16; j+=8) {
        for (i=0; i<nWidth-16; i+=8) {
            lpRec = lpBuffer + i + j * nWidth;
            nActSrc = CalcActivitybyRect(lpRec, nWidth, 0, 8, 8); // Активность левого блока
            nActSrc2 = CalcActivitybyRect(lpRec+8, nWidth, 0, 8, 8); // Активность правого блока
            nActSrc = (nActSrc + nActSrc2)/2; // Среднее значений активности
            nDiff = 0;
            for(int jj=0; jj<8; jj++){
                nDiff += abs(lpRec[7+jj*nWidth] - lpRec[8+jj*nWidth]);
                // Разница значений яркости на границе
            }
            nDiff/= 8;
        }
    }
}

```

```
        dBlock += (double)nDiff/(double)(nActSrc+1);
        nMB++;
    }
}
return (double)dBlock/(double)nMB;
}
```

## 6 Информационные справочные документы

- [1] SMPTE 170M, "SMPTE Standard for Television – Composite Analog Video Signal – NTSC for Studio Applications", Society of Motion Picture and Television Engineers.
- [2] ITU-T Recommendation J.143 – User requirements for objective perceptual video quality measurements in digital cable television.

## Приложение С

### Модель С: Метод с использованием эталонного сигнала с уменьшенной полосой частот, разработанный NTIA

#### 1 Базовая информация

В период 2003–2004 годов Национальное управление по связи и информации США (NTIA) разработало две модели оценки качества изображения (VQM) с использованием эталонного сигнала с уменьшенной полосой частот (RR), приблизительно 12–14 кбит/с, для дискретизованного изображения, соответствующего Рекомендации МСЭ-R ВТ.601 [1]. Эти модели получили названия "узкополосная VQM" и "скоростная узкополосная VQM". Скоростная узкополосная VQM представляла собой эффективную в вычислительном отношении версию узкополосной VQM. Скоростная узкополосная VQM примерно в четыре раза быстрее, поскольку она осуществляет извлечение пространственных свойств из видеокadres, которые заранее усреднены, а не извлечение пространственных свойств непосредственно из видеокadres ITU-R ВТ.601. Дополнительная экономия вычислительных ресурсов в случае скоростной узкополосной VQM достигается благодаря обработке временных свойств информации (т. е. движения) на основании случайной подвыборки пикселей в канале яркости Y вместо использования всех пикселей во всех трех видеоканалах (Y, Cb и Cr). Обе VQM на протяжении нескольких лет входят в состав наших программных инструментов VQM и могут без ограничений применяться как для коммерческих, так для некоммерческих приложений. Двоичные исполняемые версии этих инструментов VQM и их соответствующие исходные коды доступны для загрузки [2].

Поскольку NTIA намеревалось представить как узкополосную VQM, так и скоростную узкополосную VQM на испытаниях ТВ с эталонным сигналом с уменьшенной полосой частот (RRTV) для независимой оценки Группой экспертов по качеству видеосигнала (VQEG), NTIA представило их в различных категориях скорости, несмотря на идентичные требования к скорости RR. NTIA представило узкополосную VQM в категории 256k, а скоростную узкополосную VQM – в категории 80k, поскольку ожидалось, что показатели работы узкополосной VQM будут выше показателей скоростной узкополосной VQM. Обе VQM использовали алгоритм калибровки RR, представленный в Рекомендации МСЭ-T J.244 [3]. Для этого алгоритма калибровки требуется примерно от 22 до 24 кбит/с ширины полосы RR для получения оценок временной задержки, пространственного сдвига, пространственной градации, усиления и сдвига уровня.

Интересный результат оценочных испытаний VQEG RRTV [4] заключается в том, что показатели работы скоростной узкополосной VQM превысили показатели узкополосной VQM в испытании как для формата 525, так и 625 строк. Это вызывает интерес, так как означает, что извлечение пространственных свойств из усредненных кадров эффективнее, чем их извлечение из неусредненных кадров. Далее предстоит изучить, справедлив ли этот результат для других наборов

данных. В настоящее время NTIA не видит оснований стандартизировать обе модели, ввиду чего в настоящем Приложении описывается только скоростная узкополосная VQM.

## 2 Введение

В настоящем Приложении приводятся описание и машинная программа для скоростной узкополосной VQM NTIA. Скоростная узкополосная VQM NTIA использует методы, аналогичные методам общей VQM NTIA, описание которой содержится как в Рекомендации МСЭ-T J.144 [5], так и в Рекомендации МСЭ-R ВТ.1683 [6]. Скоростная узкополосная VQM использует свойства RR со значительно меньшей шириной полосы, чем используемая в общей VQM NTIA, но процесс извлечения свойств и сравнения аналогичен для обеих. Для дискретизованного изображения, соответствующего Рекомендации МСЭ-R ВТ.601 [1], скоростная узкополосная VQM использует свойства RR, для которых требуется примерно 12–14 кбит/с ширины полосы передачи. В настоящем Приложении описывается только скоростная узкополосная VQM, поскольку ее являющиеся дополняющими алгоритмы калибровки с узкой полосой частот подробно описаны в Рекомендации МСЭ-T J.244 [3]. Вместе с тем для полноты машинная программа в настоящем Приложении включает как скоростную узкополосную VQM, так и связанные с ней алгоритмы калибровки с узкой полосой частот. Машинная программа содержит также примеры функций квантования для свойств, используемых при калибровке с узкой полосой частот (эти функции квантования не являются частью Рекомендации МСЭ-T J.244).

## 3 Описание скоростной узкополосной VQM

### 3.1 Обзор VQM

Описание VQM охватывает три основных области:

- 1) узкополосные свойства, которые извлекаются из исходного и обработанного видеопотоков;
- 2) параметры, являющиеся результатом сравнения исходных и обработанных потоков свойств;
- 3) расчет VQM, в котором сочетаются различные параметры, каждый из которых измеряет отдельный аспект качества изображения.

В том что касается технических подробностей, в настоящем описании используются имеющиеся справочные документы.

### 3.2 Описание свойств

#### 3.2.1 Обзор свойств

Скоростная узкополосная VQM использует три типа свойств: цветовые, пространственные и временные. Каждое из этих свойств выражает в количественном отношении воспринимаемые искажения в соответствующей области. Подпрограмма "model\_fastlowbw\_features" в приведенном примере программы обеспечивает полное математическое описание характеристик, используемых скоростной узкополосной VQM.

#### 3.2.2 Цветовые свойства

Цветовые свойства – это те же свойства  $f_{COHER\_COLOR}$ , что и используемые общей VQM NTIA. Эти свойства подробно описываются в Приложении D.7.3 к Рекомендации МСЭ-T J.144. Свойства  $f_{COHER\_COLOR}$  обеспечивают двухмерное векторное измерение количества информации о цветности по синему и красному цвету (CB, CR) в каждой S-T зоне. Таким образом, свойства  $f_{COHER\_COLOR}$  чувствительны к искажению цвета. Свойства  $f_{COHER\_COLOR}$  скоростной узкополосной VQM NTIA извлекаются из пространственно-временной (S-T) зоны размером 30 вертикальных линий на 30 горизонтальных пикселей на 1 с времени (т. е.  $30 \times 30 \times 1$  с), в то время как в общей VQM NTIA используются размер зоны S-T  $8 \times 8 \times 1$  кадр.



### 3.2.3 Пространственные свойства

Пространственные свойства – это те же свойства  $f_{SI13}$  и  $f_{HV13}$ , что и используемые общей VQM NTIA. Эти свойства подробно описываются в Приложении D.7.2.2 к Рекомендации МСЭ-T J.144. Свойства  $f_{SI13}$  и  $f_{HV13}$  обеспечивают измерение количества и углового распределения пространственных градиентов в пространственно-временных (S-T) подзонах изображений в канале яркости (Y). Таким образом, свойства  $f_{SI13}$  и  $f_{HV13}$  чувствительны к пространственным искажениям, таким как размытость и блокирование.

Свойства  $f_{SI13}$  и  $f_{HV13}$  скоростной узкополосной VQM NTIA извлекаются из пространственно-временной (S-T) зоны размером 30 вертикальных линий на 30 горизонтальных пикселей на 1 секунду времени (т. е.  $30 \times 30 \times 1$  с), тогда как в общей VQM NTIA используется размер зоны S-T  $8 \times 8 \times 0,2$  с. Наряду с этим для экономии вычислений изображения 1 с в канале яркости Y заранее усредняются по времени до применения двумерных фильтров коррекции контуров  $13 \times 13$ , описываемых в Приложении D.7.2.1 к Рекомендации МСЭ-T J.144.

Одна дополнительное пространственное свойство извлекается из 1 с предварительно усредненных изображений в канале яркости (Y). Этим свойством является *средний* уровень яркости каждой зоны S-T  $30 \times 30 \times 1$  с (здесь обозначается как  $f_{MEAN}$ ). Назначение свойства  $f_{MEAN}$  – обеспечить функцию взвешивания восприятия уровня яркости для потери пространственной информации (SI), измеряемой с помощью свойств  $f_{SI13}$  и  $f_{HV13}$ . Это описывается в разделе, посвященном описанию параметров.

### 3.2.4 Временные свойства

Из набора цветовых и пространственных свойств, описанных выше, можно получить существенные оценки воспринимаемого качества изображения. Вместе с тем, поскольку S-T зоны, из которых извлекаются эти свойства, охватывают большое число видеокадров (т. е. 1 с видеокадров), они, как правило, нечувствительны к кратким временным искажениям изображения. Такие искажения могут быть следствием шума или ошибок цифровой передачи; и, несмотря на краткость, они могут оказывать существенное влияние на воспринимаемое качество изображения. Таким образом, для количественного выражения восприятия временных искажений используется свойство RR, основанное на времени. С помощью этого свойства измеряется абсолютная временная информация (ATI), или движение, в плоскости изображения в канале яркости Y и рассчитывается следующим образом:

$$f_{ATI} = rms\{rand5\%|Y(t) - Y(t - 0.2s)|\}.$$

Для эффективности вычислений из Y произвольно берется подвыборка, содержащая только 5% пикселей изображения (здесь представленная функцией *rand5%*). Изображение Y, по которому сделана подвыборка в момент времени  $t - 0,2$  с вычитается из изображения Y, по которому идентичным образом сделана подвыборка в момент времени  $t$ , и средняя квадратичная ошибка (*rms*) результата используется как мера ATI. Используя условные обозначения, приведенные в Приложении D.8 к Рекомендации МСЭ-T J.144, это также обозначается следующим образом:

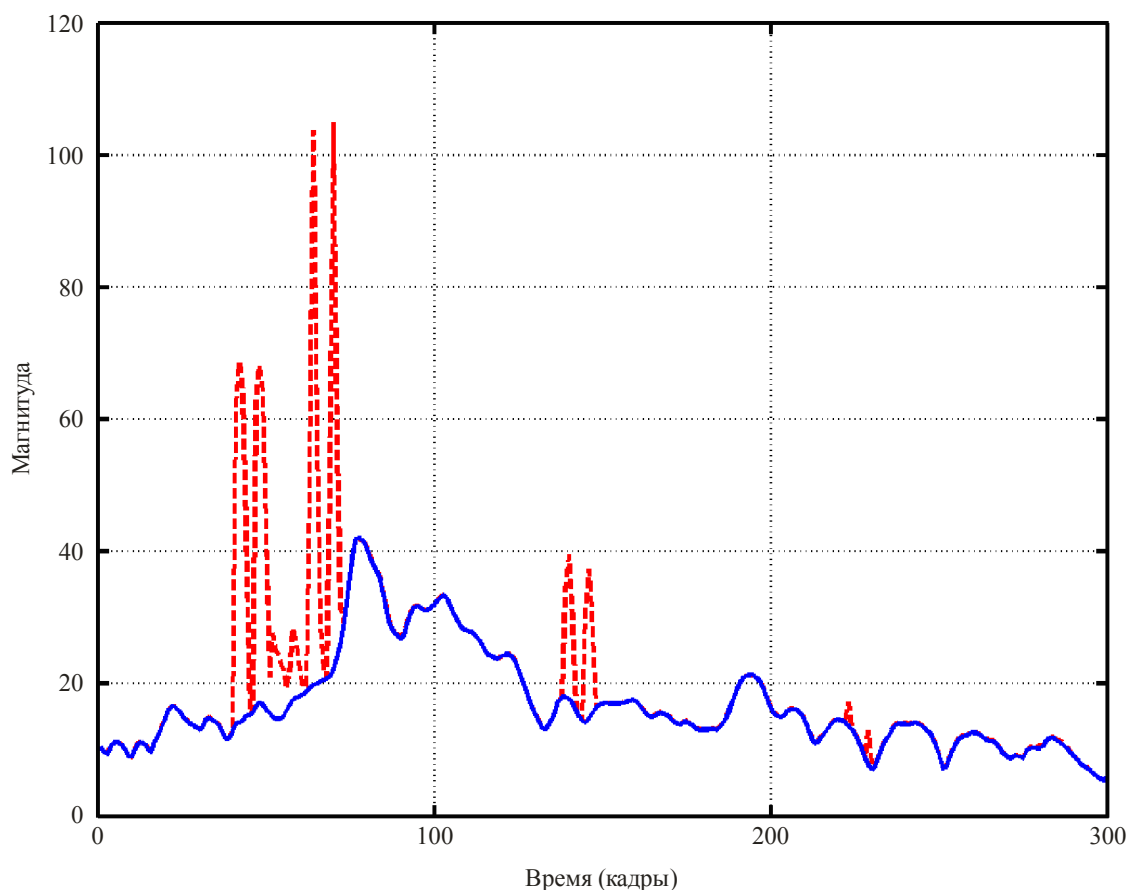
$$f_{ATI} \cong Y\_rand5\%\_ati0.2s\_rms.$$

Свойство  $f_{ATI}$  чувствительно к временным искажениям. Для изображения 30 кадр/с период 0,2 с представляют 6 видеокадров, тогда как для изображения 25 кадр/с период 0,2 с представляют собой 5 видеокадров. Вычитание изображений, разнесенных на 0,2 с, делает это свойство нечувствительным в режиме реального времени для видеосистем формата 30 кадр/с и 25 кадр/с с частотой кадров не менее 5 кадр/с. Аспекты качества таких видеосистем с низкой частотой кадров, которые являются общими в мультимедийных приложениях, в достаточной мере описываются свойствами  $f_{SI13}$ ,  $f_{HV13}$  и  $f_{COHER\_COLOR}$ . Разнесение на 0,2 с также более приближено к пиковой динамической реакции зрения человека по сравнению с различением двух изображений, отстоящих друг от друга во времени на один кадр.

На рисунке 27 представлен пример графика свойства  $f_{ATI}$  для видеосцены источника (сплошная синяя линия) и обработанной видеосцены (пунктирная красная линия) от цифровой видеосистемы со случайными пакетными ошибками в канале цифровой передачи. Случайные ошибки в обработанном изображении создают зубцы в свойстве  $f_{ATI}$ . Для передачи свойства  $f_{ATI}$  требуется чрезвычайно малая ширина полосы, поскольку для этого необходимы лишь 30 выборок в секунду в случае изображения 30 кадр/с. Другие виды аддитивного шума в обработанном изображении, который может производиться аналоговой видеосистемой, будут отображаться как положительный сдвиг DC во временной диаграмме обработанного потока свойств относительно исходного потока свойств. Системы кодирования видеосигнала, устраняющие шум, вызовут отрицательный сдвиг DC.

До извлечения параметра случайных ошибок из потоков свойств  $f_{ATI}$ , показанных на рисунке 27, полезно увеличить ширину зубцов движения (красных зубцов на рисунке 27). Это обосновывается тем, что короткие зубцы движения от случайных ошибок не обеспечивают адекватного представления воздействия этих типов ошибок на восприятие. Один из методов увеличения ширины зубцов движения является применение максимального фильтра как к исходному, так и к обработанному потоку свойств до расчета функции параметра ошибок между двумя формами сигнала. Для параметра ошибок на базе  $f_{ATI}$  использовался максимальный фильтр шириной 7 пунктов (который здесь обозначается как функция  $max7pt$ ), вырабатывающий в каждом кадре выходную выборку, которая является максимумом собственной и трех ближайших примыкающих выборок с каждой стороны (т. е. предыдущие и последующие по времени выборки).

РИСУНОК 27

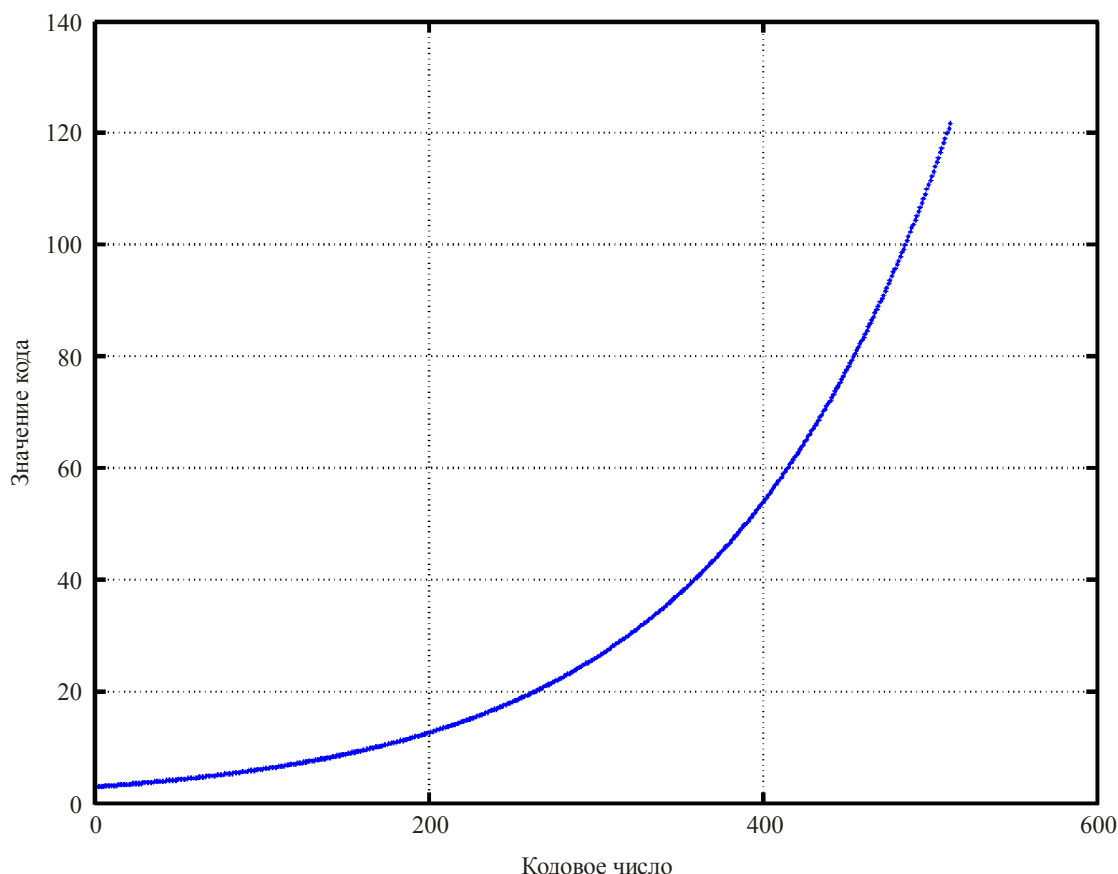
Пример временной истории характеристики  $f_{ATI}$ 

### 3.2.5 Квантование характеристик

Для свойств  $Y_{MEAN}$ ,  $f_{SI13}$ ,  $f_{HV13}$  и  $f_{COHER\_COLOR}$  достаточно квантования до 9 битов точности, тогда как свойство  $f_{ATI}$  следует квантовать до 10 битов. Для сведения к минимуму воздействия на расчеты параметров качества изображения следует использовать нелинейный квантователь, в котором погрешность квантователя пропорциональна величине квантуемого сигнала. Очень малые значения обычно квантуются до некоего предельного значения, ниже которого полезная информация для оценки качества отсутствует. Такой вариант квантователя сводит к минимуму погрешности в соответствующих расчетах параметров, поскольку эти расчеты обычно базируются на коэффициенте ошибок или логарифмическом отношении обработанного и исходного потоков свойств (см. раздел, посвященный описанию параметров, ниже).

На рисунке 28 представлен график 9-битного нелинейного квантователя, используемого для исходного свойства  $f_{SI13}$ . Подпрограмма "model\_lowbw\_compression" в приведенном примере программы обеспечивает полное математическое описание рекомендуемых квантователей, используемых скоростной узкополосной VQM. Если свойства выходят за пределы охвата рекомендуемых квантователей у нижнего или верхнего предела (что крайне маловероятно), то S-T параметры, извлеченные из этих свойств, устанавливаются равными нулю и поэтому не влияют на VQM в целом.

РИСУНОК 28

Нелинейный 9-битный квантизатор для характеристики  $f_{SI13}$ 

### 3.3 Описание параметров

#### 3.3.1 Краткий обзор параметров

Расчет параметров, характеризующих различные аспекты восприятия качества изображения, предусматривает ряд шагов, в том числе:

- применение порога восприятия к свойствам, извлеченным из каждой S-T подзоны;
- расчет функции ошибок между обработанными свойствами и соответствующими исходными свойствами;
- объединение полученных ошибок по пространству и времени.

В Приложении D.8 к Рекомендации МСЭ-T J.144 содержится подробное описание этих методов и соответствующая им математическая запись названий параметров, которые также будут применяться здесь. Подпрограмма "model\_fastlowbw\_parameters" в приведенном примере программы обеспечивает полное математическое описание параметров, используемых скоростной узкополосной VQM. Для целей упрощения в описании параметров в настоящем разделе не учитываются последствия квантования свойств (например, обработка значений свойств, которые могут попадать за пределы рекомендуемых диапазонов квантования).

#### 3.3.2 Новые методы

В настоящем разделе кратко излагаются новые методы обеспечения более точной корреляции объективных и субъективных параметров на базе свойств RR с очень малыми значениями ширины полосы передачи, такими как применяемые для скоростной узкополосной VQM NTIA (т. е. новые методы, не описанные в Рекомендации МСЭ-T J.144). Следует отметить, что не было обеспечено улучшение базовой формы функций ошибок параметров, приведенной в Приложении D.8.2.1 к Рекомендации МСЭ-T J.144. Двумя функциями ошибок, которые последовательно дают наилучшие результаты по параметрам (для пространственных и временных параметров), являются логарифмическое отношение  $\{\log_{10} [f_p(s,t) / f_o(s,t)]\}$  и коэффициент ошибок  $\{[f_p(s,t) - f_o(s,t)] / f_o(s,t)\}$ , где  $f_p(s,t)$  и  $f_o(s,t)$  – обработанное свойство и соответствующая исходное свойство, извлеченные из S-T зоны с пространственными координатами  $s$  и временными координатами  $t$  соответственно. Ошибки следует подразделять на усиление и потери, поскольку люди по-разному реагируют на аддитивные (например, блочность) и субтрактивные (например, размывание) искажения. Применение более низкого порога восприятия к свойствам для применения этих двух функций ошибок деление на ноль предотвращает.

После расчета S-T параметров с использованием одной из функций ошибок S-T параметры должны быть объединены по пространству и времени для получения значения параметра видеофрагмента. Это объединение ошибок может происходить несколькими этапами (например, по пространству, а затем по времени). Один из новых методов объединения, используемых в скоростной узкополосной VQM носит название объединения ошибок макроблоками (MB). При объединении ошибок методом MB группируется несколько смежных S-T подзон и к этому множеству применяется функция объединения ошибок. Например, функция, обозначаемая как "MB(3,3,2)max", будет применять функцию max к значениям параметров из каждой группы 18 S-T подзон, формат объединения которых – 3 по вертикали, 3 по горизонтали и 2 временных. Для  $32 \times 32 \times 1$  с S-T подзон свойств  $f_{SI13}$ ,  $f_{HV13}$  и  $f_{COHER\_COLOR}$ , описанных выше, каждая зона MB(3,3,2) будет охватывать часть видеопотока, занимающего 96 вертикальных линий на 96 горизонтальных пикселей на 2 секунды времени. Объединение ошибок MB было признано полезным для отслеживания воздействия на восприятие искажений, локализованных в пространстве и времени. Такие локализованные искажения зачастую влияют на процесс принятия решений о качестве. Объединение ошибок MB может также осуществляться как процесс фильтрации, при котором вместо получения единого значения на выходе для каждого MB каждый S-T образец заменяется его значением, фильтрованным MB, где MB центруется на выборке S-T. Это носит название перекрывающегося объединения ошибок MB (OMB).

Вторым методом объединения ошибок является обобщенная сумма Минковского ( $P, R$ ), определяемая как:

$$Minkowski(P, R) = R \sqrt[\frac{R}{P}]{\frac{1}{N} \sum_{i=1}^N |v_i|^P}.$$

Здесь  $v_i$  представляет значения параметров, которые включены в сумму. Так, эта сумма может включать все значения параметров на данный момент времени (пространственное объединение) или же может применяться к макроблокам, описанным выше. Сумма Минковского, где степень  $P$  равна корню  $R$ , использовалась многими разработчиками метрик качества изображения для объединения ошибок. Обобщенная сумма Минковского, где  $P \neq R$ , обеспечивает дополнительную гибкость для линеаризации реакции отдельных параметров на изменения в воспринимаемом качестве. Это необходимый шаг перед комбинированием множества параметров в единой оценке воспринимаемого качества изображения, которое осуществляется методом линейного приближения.

### 3.3.3 Параметры цвета

Из свойств  $f_{COHER\_COLOR}$  извлекаются два параметра. Один из этих параметров,  $color\_extreme$ , измеряет предельные искажения цвета, которые могут вызываться цветными блоками от ошибок передачи. Другой параметр,  $color\_spread$ , дает представление о дисперсии или разброс ошибок цвета. Вместе использования эвклидовой меры расстояния для количественного выражения искажений (как в Приложении D.8.2.2 к Рекомендации МСЭ-Т J.144) оба этих параметра используют корень квадратный из манхэттенского расстояния. В соответствии с математической записью в Приложении D.8.2.2 к Рекомендации МСЭ-Т J.144, где  $f_p(s, t)$  и  $f_o(s, t)$  представляют двухмерное свойство  $f_{COHER\_COLOR}$ , извлеченное из S-T зоны обработанного и исходного видеопотоков, эта функция сравнения свойств выражается следующим образом:

$$sqrtmanhat(s, t) = \sqrt{\sum_{C_B, C_R} \left| \frac{f_p(s, t) - f_o(s, t)}{-p} \right|}.$$

Манхэттенская мера расстояния представляется лучше эвклидовой меры расстояния, а функция квадратного корня необходима для линеаризации реакции параметра на изменения качества. Согласно математической записи в Приложении D.8 к Рекомендации МСЭ-Т J.144, параметры цвета выражаются следующим образом:

$$color\_extreme = color\_coher\_color\_30x30\_1s\_mean\_sqrtmanhat\_OMB(3,3,2)above99\%\_Minkowski(0.5,1)$$

$$color\_spread = color\_coher\_color\_30x30\_1s\_mean\_sqrtmanhat\_OMB(3,3,2)Minkowski(2,4)\_90\%.$$

Затем комбинированный параметр цвета ( $color\_comb$ ), который содержит оптимальную комбинацию параметров  $color\_extreme$  и  $color\_spread$ , рассчитывается как:

$$color\_comb = 0.691686 * color\_extreme - 0.617958 * color\_spread.$$

Далее этот имеющий положительное значение параметр  $color\_comb$  отсекается на нижнем конце, что математически представляется следующим образом (следуя математической записи в Приложении D.8.5 к Рекомендации МСЭ-Т J.144):

$$color\_comb = color\_comb\_clip\_0.114.$$

Этот параметр  $color\_comb$  включается в линейную комбинацию для расчета VQM.

### 3.3.4 Пространственные параметры

По свойству  $f_{SI13}$  рассчитываются два пространственных параметра: один измеряет потерю пространственной информации ( $si\_loss$ ), а другой измеряет усиление пространственной информации ( $si\_gain$ ). Согласно математической записи в Приложении D.8 к Рекомендации МСЭ-T J.144, эти параметры выражаются следующим образом:

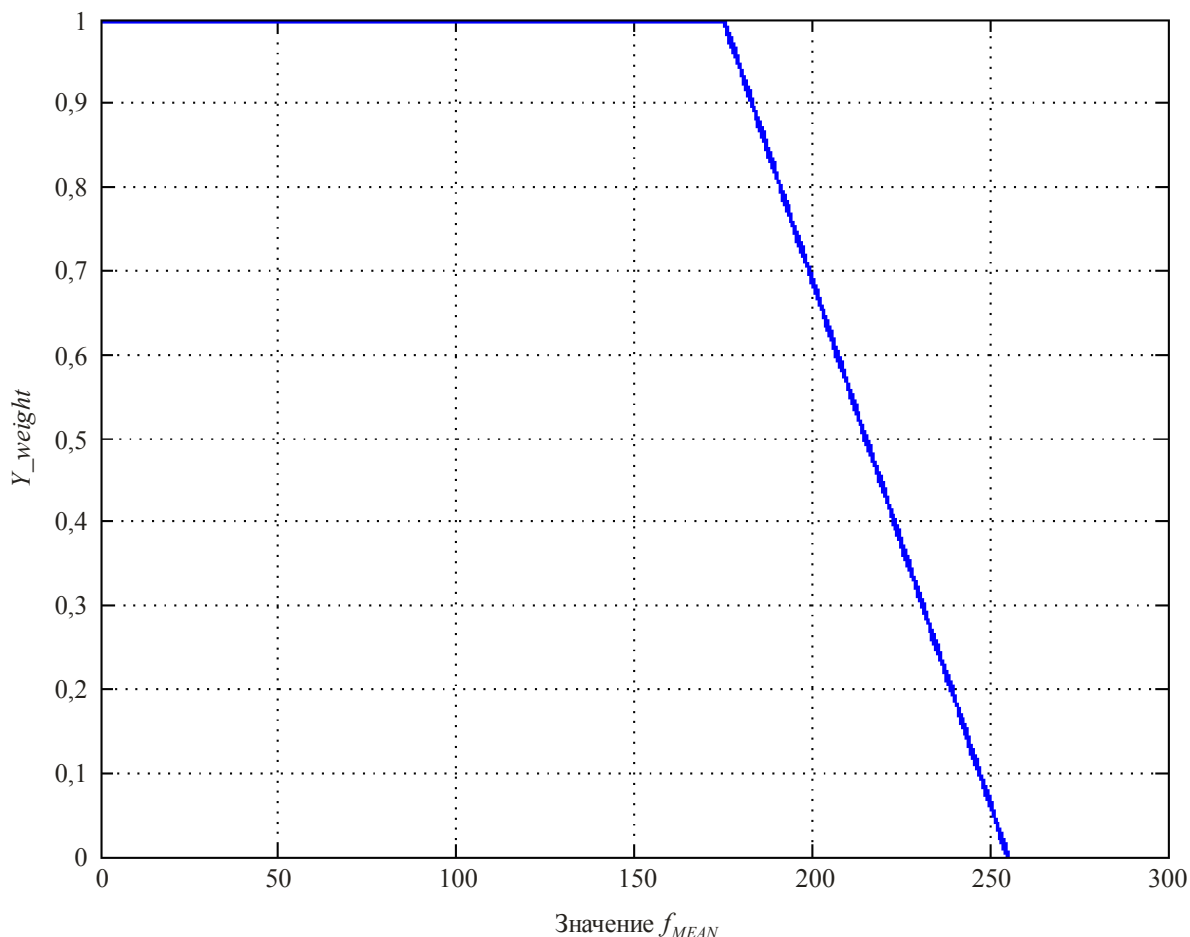
$$si\_loss = avg1s\_Y\_si13\_30x30\_std\_3\_ratio\_loss\_OMB(3,3,2)Minkowski(1,2)\_Minkowski(1.5,2.5)\_clip\_0.12$$

$$si\_gain = avg1s\_Y\_si13\_30x30\_std\_3\_log\_gain\_clip\_0.1\_above95\%tail\_Minkowski(1.5,2).$$

По мере увеличения среднего уровня яркости (Y) S-T подзоны (измеряемого на основании свойства  $f_{MEAN}$ ) уменьшается способность воспринимать изменения пространственных деталей (например, размытия, измеряемого  $si\_loss$ ). Это можно учесть, введя функцию взвешивания ( $Y\_weight$ ), как показано на рисунке 29, применяемую к значениям  $si\_loss$  из каждой S-T подзоны (т. е. значений  $si\_loss$  после выполнения функции сравнения коэффициента потери по каждой S-T подзоне, но до функций пространственного и временного уплотнения). Функция взвешивания  $Y\_weight$  равна единице (т. е. полное взвешивание), пока не достигается средний уровень яркости 175, а затем она линейно уменьшается до нуля, тогда как значения яркости возрастают со 175 до 255. Эта промежуточная коррекция применяется только к значениям  $si\_loss$ , но не к значениям  $si\_gain$ .

РИСУНОК 29

Функция взвешивания  $Y\_weight$  для модификации S-T параметров  $si\_loss$



ВТ.1885-29

По свойству  $f_{HV13}$  рассчитываются два пространственных параметра: один измеряет потерю относительной горизонтальной и вертикальной (HV) информации ( $hv\_loss$ ), а другой измеряет

усиление ( $hv\_gain$ ). Согласно математической записи в Приложении D.8 к Рекомендации МСЭ-T J.144, эти параметры выражаются следующим образом:

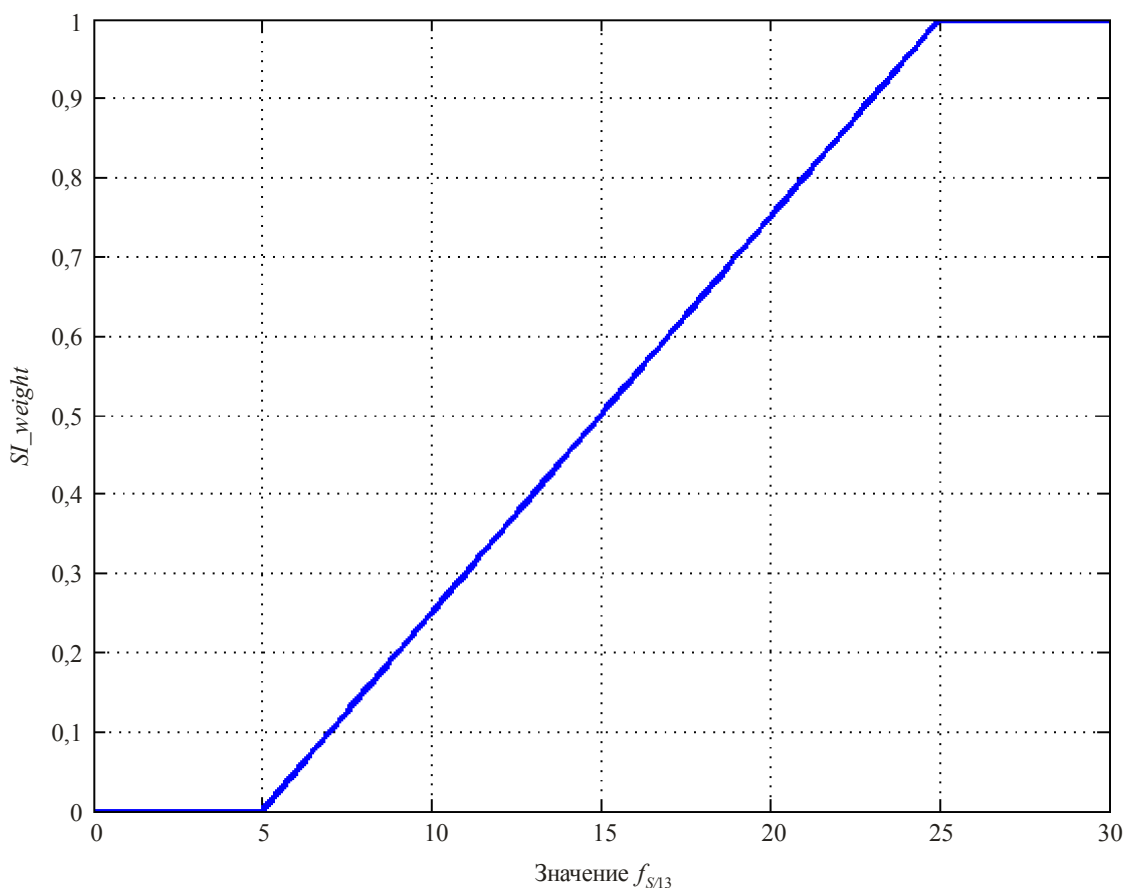
$$hv\_loss = \text{avg1s\_Y\_hv13\_angle0.225\_rmin20\_30x30\_mean\_4\_ratio\_loss\_...} \\ \text{OMB(3,3,2)below1\%\_Minkowski(1,1.5)\_clip\_0.08}$$

$$hv\_gain = \text{avg1s\_Y\_hv13\_angle0.225\_rmin20\_30x30\_mean\_4\_log\_gain\_...} \\ \text{clip\_0.06\_OMB(3,3,2)above99\%tail\_Minkowski(1.5,3)}$$

В приведенных выше уравнениях не показано, что функция  $Y\_weight$  на рисунке 29 также применяется к значениям  $hv\_loss$  и  $hv\_gain$  из каждой S-T подзоны до функций пространственного и временного уплотнения (после расчета  $ratio\_loss$  и  $log\_gain$ , соответственно). К значениям  $hv\_loss$  из каждой S-T подзоны применяется дополнительная функция взвешивания ( $SI\_weight$ , показанная на рисунке 30). Это необходимо для снижения чувствительности  $hv\_loss$  для S-T зон, которые располагают весьма небольшим объемом пространственной информации (т. е. низкие исходные значения свойства  $f_{SI3}$ ).

РИСУНОК 30

Функция взвешивания  $SI\_weight$  для модификации S-T параметров  $hv\_loss$



ВТ.1885-30

Параметры пространственного искажения могут быть сжаты (т. е. чрезмерные выбросы за пределы квалификационных данных ограничиваются или сжимаются) с использованием таких функций, как функция сжатия VQM, приведенная в разделе расчетов VQM.

### 3.3.5 Временные параметры

По свойству  $f_{ATI}$  рассчитываются два временных параметра: один измеряет добавленный случайный шум ( $ati\_noise$ ), а другой измеряет нарушения движения, вызванные ошибками передачи ( $ati\_error$ ).

Согласно математической записи в Приложении D.8 к Рекомендации МСЭ-T J.144, эти параметры выражаются следующим образом:

$$ati\_noise = Y\_rand5\%\_ati0.2s\_rms\_5\_ratio\_gain\_between25\%50\%$$

$$ati\_error = Y\_rand5\%\_ati0.2s\_rms\_max7pt\_12\_ratio\_gain\_above90\%.$$

Для усиления устойчивости параметров  $ati\_noise$  и  $ati\_error$  к временным рассогласованиям, эти параметры рассчитываются для всех временных согласований обработанного изображения, которые находятся в пределах  $\pm 0,4$  с от наилучшего оценочного временного согласования с исходным изображением, а затем выбирается минимальное значение параметра.

### 3.4 Расчет VQM

Аналогично общей VQM NTIA в Приложении D к Рекомендации МСЭ-T J.144, в расчете скоростной узкополосной VQM линейным образом сочетаются два параметра из свойства  $f_{HV13}$  ( $hv\_loss$  и  $hv\_gain$ ), два параметра из свойства  $f_{SI13}$  ( $si\_loss$  и  $si\_gain$ ) и два параметра из свойства  $f_{COHER\_COLOR}$  (но эти два параметра объединены в единый параметр искажения цвета  $color\_comb$ ). Единственный параметр шума, использовавшийся в общей VQM NTIA, был заменен двумя параметрами на основе узкополосного свойства  $f_{ATI}$ , описываемого здесь ( $ati\_noise$  и  $ati\_error$ ).

Таким образом,  $VQM_{FLB}$  (сокращенное название скоростной узкополосной VQM) состоит из линейной комбинации восьми параметров.  $VQM_{FLB}$  задается как:

$$VQM_{FLB} = \{ 0,38317338378290 * hv\_loss + 0,37313218013131 * hv\_gain + \\ 0,58033514546526 * si\_loss + 0,95845512360511 * si\_gain + \\ 1,07581708014998 * color\_comb + \\ 0,17693274495002 * ati\_noise + 0,02535903906351 * ati\_error \}.$$

Общее значение VQM (после сложения вкладов всех параметров) отсекается на нижнем пороге 0,0, с тем чтобы не допустить отрицательных значений VQM. В заключение к значениям VQM выше 1,0 применяется функция сжатия, которая допускает превышение максимум в 50%, с тем чтобы ограничить значения VQM для чрезмерно искаженного изображения, выходящего за пределы квалификационных данных.

Если  $VQM_{FLB} > 1,0$ , то  $VQM_{FLB} = (1 + c) * VQM_{FLB} / (c + VQM_{FLB})$ , где  $c = 0,5$ .

Рассчитанное описанным выше образом  $VQM_{FLB}$  будет иметь значения больше или равные нулю при номинальном максимальном значении, равном единице.  $VQM_{FLB}$  иногда может превышать единицу при крайне искаженных видеосценах.

Для усиления устойчивости  $VQM_{FLB}$  к пространственным рассогласованиям,  $VQM_{FLB}$  рассчитывается для всех пространственных согласований обработанного изображения, находящихся в пределах плюс/минус 1 пиксель от наилучшего оценочного пространственного согласования с исходным видео, а затем выбирается минимальное значение  $VQM_{FLB}$ .

## 4 Справочные документы

- [1] Рекомендация МСЭ-R ВТ.601-6 (01/07) – Студийные параметры кодирования цифрового телевидения для стандартного 4:3 и широкоэкранный 16:9 форматов.
- [2] Video Quality Model (VQM) Software Tools – Binary executables and source code, доступна на сайте Национального управления по связи и информации (NTIA) по адресу: [http://www.its.blrdoc.gov/n3/video/VQM\\_software.php](http://www.its.blrdoc.gov/n3/video/VQM_software.php).
- [3] ITU-T Recommendation J.244 (04/08) – Full reference and reduced reference calibration methods for video transmission systems with constant misalignment of spatial and temporal domains with constant gain and offset.



- [4] VQEG Final Report of MM Phase I Validation Test (2008), "Final report from the Video Quality Experts Group on the validation of objective models of multimedia quality assessment, phase I", Video Quality Experts Group (VQEG), <http://www.its.bldrdoc.gov/vqeg/projects/multimedia>, 9-я Исследовательская комиссия МСЭ-Т TD923, исследовательский период 2005–2008 годов.
- [5] ITU-T Recommendation J.144 (03/04) – Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference.
- [6] Recommendation ITU-R ВТ.1683 (06/04) – Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference.

## 5 Пример машинной программы для реализации скоростной узкополосной VQM

Данный пример машинной программы представлен в помощь пользователям для надлежащей реализации скоростной узкополосной VQM. Для примера приведена машинная программа MATLAB®, но можно применять любой программный код, который воспроизводит приведенные здесь результаты. В каждом подразделе раздела 5 содержится код MATLAB для функции, обозначенной в заголовке раздела (за исключением содержания подраздела 5.1, отсылающего к файлу под названием "fastlowbw\_ref.m"). Для получения информации о том, как вызвать подпрограмму, следует выполнить fastlowbw\_ref без аргументов. Данная программа обладает гибкостью для применения модели на коротком видеофрагменте (5–15 с) в более продолжительной видеопоследовательности (например, в одноминутной последовательности). Это осуществляется путем сдвига короткого видеофрагмента на одну секунду и повторного расчета модели для каждого временного сдвига. Эта функциональная возможность ниже не показана, но комментарии в программе и возвращенные аргументы от "model\_fastlowbw\_parameters.m" будут относиться к этой возможности. Данная скрытая возможность может быть полезна для применения системы контроля качества изображения в процессе эксплуатации.

Во время обработки векторов выборочного контроля (видеофрагменты) с помощью машинной программы скоростной узкополосной VQM (функция "fastlowbw\_ref.m"), создаются текстовые файлы, содержащие результаты калибровки и расчетов на модели. Для следующего примера вызовов функций MATLAB, следует получить файлы выходных данных, аналогичные приведенным ниже (ввиду случайного характера используемых скоростной узкополосной VQM процессов результаты могут немного отличаться от представленных здесь):



5.Reference  
Code.doc

## Дополнение

### Анализ ошибок передачи

Валидационные испытания VQEG для проекта RRNR-TV включали форматы 525-строк (NTSC) и 625-строк (PAL). В каждом эксперименте использовались 12 последовательностей источника и 156 обработанных видеопоследовательностей (PVS). Из этих 156 PVS 40 последовательностей содержали ошибки передачи, а 116 содержали только ошибки кодирования. В таблицах 9 и 10 приведены RMSE и OR для PVS с ошибками передачи. Следует отметить, что RMSE и OR рассчитывались с использованием линий регрессии, полученных из всего массива данных. Другими словами, линии регрессии рассчитывались с использованием всего массива данных. Затем рассчитывались RMSE и OR для ошибок передачи с использованием PVS с ошибками передачи.

ТАБЛИЦА 9

**RMSE и OR для валидационных испытаний RRNR-TV (формат 525 строк).**

**TE: ошибки передачи**

Формат 525	BCE		C TE		Без TE	
	RMSE	OR	RMSE	OR	RMSE	OR
Модель_A_15k	0,418	0,385	0,574	0,500	0,362	0,293
Модель_A_80k	0,423	0,378	0,582	0,475	0,366	0,293
Модель_A_256k	0,424	0,378	0,584	0,475	0,367	0,293
Модель_B_80k	0,598	0,667	0,768	0,650	0,544	0,586
Модель_B_256k	0,587	0,647	0,763	0,600	0,530	0,578
Модель_C_80k	0,465	0,513	0,557	0,550	0,440	0,405
Модель_C_256k	0,511	0,609	0,584	0,450	0,495	0,578
PSNR_NTIA	0,556	0,571	0,549	0,500	0,568	0,491

ТАБЛИЦА 10

**RMSE и OR для валидационных испытаний RRNR-TV (формат 625 строк).**

**TE: ошибки передачи**

Формат 625	BCE		C TE		Без TE	
	RMSE	OR	RMSE	OR	RMSE	OR
Модель_A_15k	0,524	0,468	0,597	0,450	0,508	0,414
Модель_A_80k	0,513	0,462	0,594	0,500	0,494	0,379
Модель_A_256k	0,516	0,468	0,593	0,500	0,499	0,379
Модель_B_80k	0,887	0,724	0,545	0,500	0,986	0,716
Модель_B_256k	0,864	0,744	0,523	0,600	0,962	0,716
Модель_C_80k	0,585	0,583	0,282	0,200	0,663	0,647
Модель_C_256k	0,657	0,590	0,292	0,175	0,747	0,638
PSNR_NTIA	0,605	0,564	0,338	0,250	0,678	0,517