

МСЭ-R

Сектор радиосвязи МСЭ

Рекомендация МСЭ-R ВТ.1907-0

(01/2012)

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

Серия ВТ

**Радиовещательная служба
(телевизионная)**



Международный
союз
электросвязи

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 Резолюции МСЭ-R 1. Формы, которые владельцам патентов следует использовать для представления патентных заявлений и деклараций о лицензировании, представлены по адресу: <http://www.itu.int/ITU-R/go/patents/en>, где также содержатся Руководящие принципы по выполнению общей патентной политики МСЭ-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)
BT	Радиовещательная служба (телевизионная)
F	Фиксированная служба
M	Подвижные службы, служба радиоопределения, любительская служба и относящиеся к ним спутниковые службы
P	Распространение радиоволн
RA	Радиоастрономия
RS	Системы дистанционного зондирования
S	Фиксированная спутниковая служба
SA	Космические применения и метеорология
SF	Совместное использование частот и координация между системами фиксированной спутниковой службы и фиксированной службы
SM	Управление использованием спектра
SNG	Спутниковый сбор новостей
TF	Передача сигналов времени и эталонных частот
V	Словарь и связанные с ним вопросы

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

Электронная публикация
Женева, 2020 г.

© ITU 2020

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

РЕКОМЕНДАЦИЯ МСЭ-R ВТ.1907-0*

Методы объективного измерения воспринимаемого качества изображения для радиовещательных применений с использованием ТВЧ при наличии полного эталонного сигнала

(2012)

Сфера применения

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

Ассамблея радиосвязи МСЭ,

учитывая,

- a) что возможность автоматически измерять качество вещательного видеосигнала уже давно признана в качестве ценного ресурса для промышленности;
- b) что в Рекомендации МСЭ-R ВТ.1683 описываются объективные методы измерения воспринимаемого качества видеоизображения для цифрового телевизионного радиовещания стандартной четкости при наличии полного эталонного сигнала;
- c) что в Рекомендации МСЭ-R ВТ.709 описываются значения параметров стандартов ТВЧ для производства и международного обмена программами, а в Рекомендации МСЭ-R ВТ.500 описываются методы субъективной оценки качества видеоизображения, включая телевидение высокой четкости;
- d) что ТВЧ все более широко используется в радиовещании;
- e) что 9-я Исследовательская комиссия МСЭ-Т, по итогам доклада о ТВЧ, присланного VQEG, разработала Рекомендацию МСЭ-Т J.341, в которой описывается объективное измерение качества видеоизображения ТВЧ при наличии полного эталонного сигнала;
- f) что объективное измерение воспринимаемого качества видеоизображения ТВЧ может дополнять методы субъективной оценки,

рекомендует,

чтобы для объективного измерения воспринимаемого качества видеоизображения в радиовещательных применениях с использованием ТВЧ при наличии полного эталонного сигнала использовалась модель объективного качества видеоизображений, приведенная в Приложениях 1, 2 и 3.

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

Приложение 1

1 Введение

В настоящей Рекомендации представлен метод измерения воспринимаемого качества видеоизображений для использования в телевидении высокой четкости (ТВЧ) без интерактивных приложений в тех случаях, когда может быть применен метод измерения с полным эталоном (FR). Эта модель была разработана для оценки субъективных показателей качества.

Метод измерения с полным эталоном может применяться, когда в точке измерения легкодоступен неискаженный эталонный видеосигнал, как, возможно, обстоит дело при измерениях на отдельном оборудовании, в лабораторной цепи или в замкнутом пространстве, таком как телевизионная радиовещательная станция. Метод оценки включает как калибровку, так и объективные оценки качества видеоизображения.

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

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

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

В проверочном испытании для каждого значения разрешения предполагались следующие условия:

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

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

1.1 Применения

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

- 1) контроль качества в источнике без перерыва связи, потенциально в режиме реального времени;
- 2) удаленный контроль качества, когда в точке измерения доступна копия источника;
- 3) измерение качества для контроля системы хранения или передачи, которая использует методы сжатия и декомпрессии видеосигнала как в течение одного прохода сигнала, так и при последовательном применении этих методов;
- 4) лабораторное тестирование видеосистем.

1.2 Ограничения

Модель оценки качества видеоизображения, описанная в настоящей Рекомендации, не может быть использована для замены субъективного тестирования. Коэффициенты корреляции между двумя тщательно разработанными и выполненными субъективными тестами (т. е. в двух разных лабораториях), обычно находятся в диапазоне от 0,95 до 0,98. Если эта Рекомендация используется для сравнения видеосистем (например, для сравнения двух кодеков), целесообразно применять количественный метод, чтобы определить точность рассматриваемой для данного конкретного контекста.

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

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

2 Ссылки

Отсутствуют.

3 Определения

3.1 Термины, определенные в других местах

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

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

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

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

3.2 Термины, определенные в настоящей Рекомендации

В данной Рекомендации определяются следующие термины:

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

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

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

4 Сокращения и акронимы

В настоящей Рекомендации используются следующие сокращения и акронимы:

ACR	Absolute Category Rating (see Recommendation ITU-R BT.500)	Оценка по абсолютным категориям (см. Рекомендацию МСЭ-R BT.500)
ACR-HR	Absolute Category Rating with Hidden Reference (see Recommendation ITU-T P.910)	Оценка по абсолютным категориям со скрытым эталоном (см. Рекомендацию МСЭ-T P.910)
AVI	Audio Video Interleave	Чередование аудио и видео
DMOS	Difference Mean Opinion Score	Разностный показатель усредненного мнения
FR	Full Reference	Полный эталонный сигнал
FRTV	Full Reference TeleVision	Телевидение с полным эталонным сигналом
HRC	Hypothetical Reference Circuit	Гипотетическая эталонная цепь
ILG	VQEG's Independent Laboratory Group	Группа "Независимая лаборатория" VQEG
MOS	Mean Opinion Score	Показатель усредненного мнения
MOSp	Mean Opinion Score, predicted	Прогнозируемый показатель усредненного мнения
NR	No (or Zero) Reference	Нет эталона (или нулевой эталон)
PSNR	Peak Signal-to-Noise Ratio	Пиковое отношение сигнал/шум
PVS	Processed Video Sequence	Обработанная видеопоследовательность
RMSE	Root Mean Square Error	Среднеквадратическая ошибка
RR	Reduced Reference	Эталон с ухудшенными характеристиками
SFR	Source Frame Rate	Частота кадров источника
SRC	Source Reference Channel or Circuit	Канал (цепь) исходного эталонного сигнала
VQEG	Video Quality Experts Group	Экспертная группа по качеству видеоизображений
YUV	Colour Space and file format	Цветовое пространство и формат файла

5 Условные обозначения

Отсутствуют.

6 Описание методологии с использованием полного эталонного сигнала

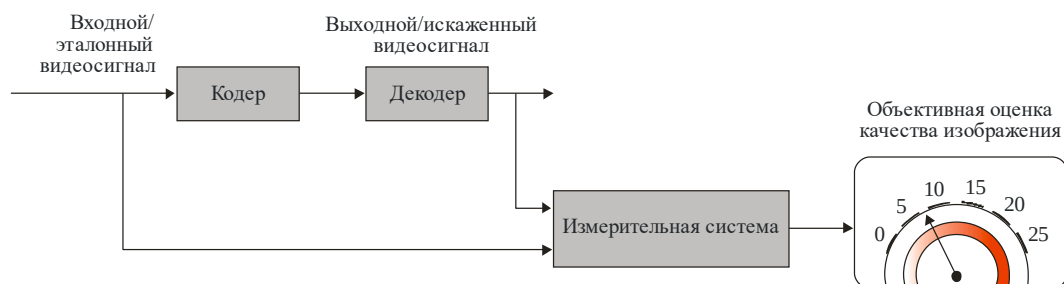
Двусторонний метод измерения с использованием полного эталона для объективного измерения воспринимаемого качества видеоизображения оценивает характеристики системы путем сравнения

между неискаженным входным (или эталонным) видеосигналом на входе системы и искаженным сигналом на выходе системы (рис. 1).

На рисунке 1 показан пример применения метода полного эталона для тестирования кодека в лаборатории.

РИСУНОК 1

Применение метода измерения воспринимаемого качества изображения при наличии полного эталона для тестирования кодека в лаборатории



ВТ.1907-01

Для сравнения между входными и выходными сигналами может потребоваться процесс временного или пространственного выравнивания, причем последнее нужно для компенсации любых вертикальных или горизонтальных смещений или подрезок изображений. Кроме того, может потребоваться коррекция для любых различий в смещении ли в коэффициентах усиления как в каналах яркости, так и в каналах цветности. Объективная оценка качества изображения затем рассчитывается, как правило, путем применения модели восприятия человеческого зрения.

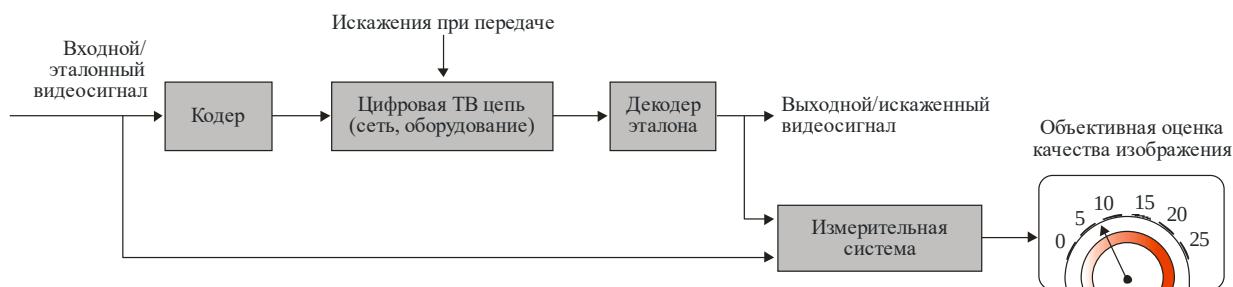
Выравнивание и регулировку усиления называют регистрацией. Этот процесс необходим, поскольку в большинстве методов, использующих полный эталон, сравниваются эталонные и обработанные изображения фактически на основе попиксельного анализа. Метрики качества видеоизображения, описанные в Приложении 2, включают методы регистрации.

Поскольку метрики качества видеоизображения, как правило, основаны на аппроксимации визуальных реакций человека, а не на оценке отдельных артефактов кодирования, они, по существу, имеют одинаковую силу для аналоговых и цифровых систем. Кроме того, они в принципе справедливы для сетей, где аналоговые и цифровые системы смешаны, или там, где последовательно включены цифровые системы сжатия.

На рисунке 2 показан пример применения метода полного эталона для проверки цепи передачи.

РИСУНОК 2

Применение метода измерения воспринимаемого качества изображения при наличии полного эталона для тестирования цепи передачи



ВТ.1907-02

В этом случае сигнал на декодер эталонного качества поступает из разных точек цепи передачи, например декодер может быть расположен в точке сети, как показано на рис. 2, или непосредственно на выходе кодера, как показано на рис. 1. Если цифровая цепь передачи сквозная, то измерение

объективной оценки качества изображения у источника эквивалентно измерению в любой последующей точке цепи.

Принято считать, что метод полного эталона обеспечивает лучшую точность измерения воспринимаемого качества изображения. Было доказано, что метод обладает возможностями тесной корреляции с субъективными оценками в соответствии с методами ACR-HR, описанными в МСЭ-Т Р.910.

7 Выводы экспертной группы по качеству видеоизображений (VQEG)

Исследования, связанные измерениями воспринимаемого качества видеоизображений, проводятся в рамках неофициальной группы, называемой экспертная группа по качеству видеоизображений (VQEG), которая отчитывается перед 9-й и 12-й Исследовательскими комиссиями МСЭ-Т и 6-й Исследовательской комиссией МСЭ-R. Недавно завершенная первая фаза тестирования телевидения высокой четкости (ТВЧ), проведенная VQEG, оценила эффективность предлагаемых алгоритмов измерения воспринимаемого качества видеоизображений при наличии полного эталонного сигнала.

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

Первичный анализ

Характеристики модели FR (модели с полным эталоном) представлены в таблице 1. Пиковое отношение сигнал/шум (PSNR) вычисляется в соответствии с МСЭ-Т J.340 и включено в этот анализ в целях сравнения. Графа "RMSE супермножества" определяет основной показатель (RMSE), рассчитанный на агрегированном супермножестве (т. е. все шесть экспериментов, отображенных на одной шкале). Графа "Наилучший общий показатель в группе" определяет количество экспериментов (от 0 до 6), для которых эта модель была либо наиболее эффективной моделью, либо статистически эквивалентна наиболее эффективной модели. Графа "Лучше, чем общее отношение PSNR" определяет количество экспериментов (от 0 до 6), для которых модель была статистически лучше, чем PSNR. Графа "Лучше, чем супермножество PSNR" указывает, является ли каждая модель статистически лучшей, чем PSNR на агрегированном супермножестве. Графа "Корреляция супермножества" определяет корреляцию Пирсона, рассчитанную на агрегированном супермножестве.

ТАБЛИЦА 1

Показатель	PSNR	SwissQual
RMSE супермножества	0,71	0,56
Наилучший общий показатель в группе	1	5
Лучше, чем общее отношение PSNR	—	4
Лучше, чем супермножество PSNR	—	Да
Корреляция супермножества	0,78	0,87

Приложение 2

Описание модели

Примечание редактора: исходный код, который включен в этот раздел, является обязательной частью настоящей Рекомендации и доступен по адресу: <http://ifatemp.itu.int/t/2009/sg9/exchange/q2/>.

Обзор модели

Данная модель прогнозирует качество видеоизображения, как оно воспринимается участниками эксперимента. Модель прогнозирования использует психо-визуальное и когнитивно-инспирированное моделирование для имитации субъективного восприятия.

Будучи основана на концепции полного эталона, данная модель сравнивает в ходе тестирования входную или высококачественную эталонную видеопоследовательность и соответствующую видеопоследовательность ухудшенного качества. Этот процесс показан на рис. 3.

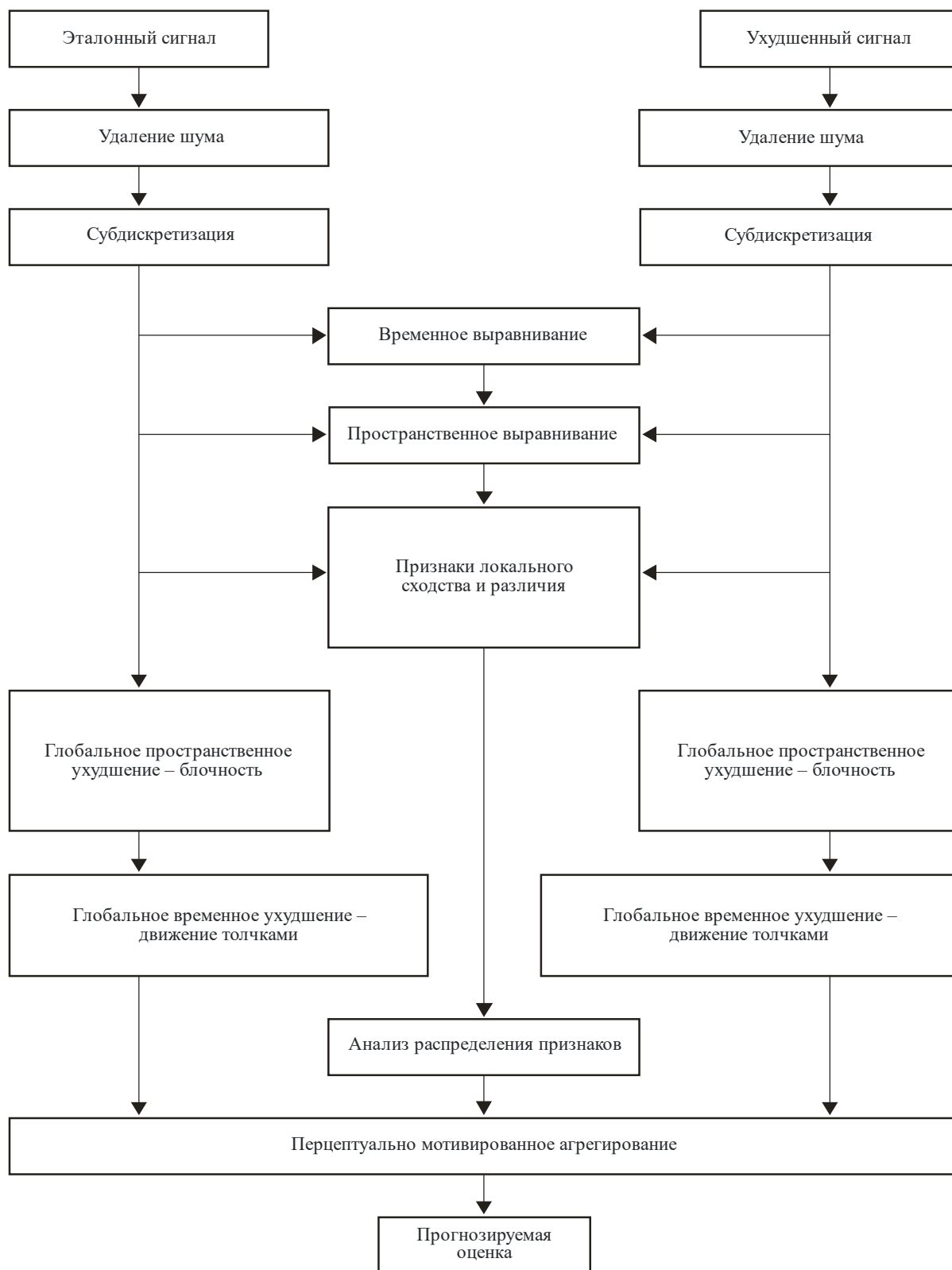
Расчет итоговой оценки основан на выполнении следующих шагов:

- 1) сначала видеопоследовательности проходят предварительную обработку. В частности, удаляется шум путем фильтрации кадров, а кадры подвергаются субдискретизации;
- 2) производится временное выравнивание кадров между эталонными и обработанными видеопоследовательностями;
- 3) производится пространственное выравнивание кадров между обработанным видеокадром и соответствующим эталонным видеокадром;
- 4) вычисляются местные пространственные характеристики качества: признаки локального сходства и различия, инспирированные зрительным восприятием;
- 5) выполняется анализ распределения признаков локального сходства и различия;
- 6) измеряется глобальное пространственное ухудшение с помощью показателя blockiness (блочности или распада изображения на квадраты);
- 7) измеряется глобальное временное ухудшение с помощью показателя jerkiness (движение толчками). Показатель движения толчками рассчитывается путем оценки локальных и глобальных значений интенсивности движения и времени отображения кадра;
- 8) определяется показатель качества на основе нелинейного агрегирования выше рассмотренных характеристик;
- 9) во избежание неправильного прогнозирования в случае относительно большого пространственного рассогласования между эталонной и обработанной видеопоследовательностями, эти шаги рассчитываются для трех различных горизонтальных и вертикальных пространственных выравниваний видеопоследовательности, и максимальный прогнозируемый показатель среди всех пространственных положений является окончательной оценкой показателя качества.

Отдельные шаги описываются более подробно в пп. 2.1–2.9. Раздел 2.10 содержит встроенный архив с исходным кодом на C++, охватывающим основные части и функции для описания совместимой реализации модели. Имена функций на C++, указанные в пп. 2.1–2.9, ссылаются на эталонный исходный код (например, п. 2.2 ссылается на `CFrameAnalysisFullRef::ContentTimeAlignment`).

РИСУНОК 3

Блок-схема этапов обработки сигналов в модели. Наверху входной сигнал – это эталонная и искаженная (или обработанная) видеопоследовательности. Различные этапы обработки дают в результате основную выходную последовательность модели, прогнозируемая оценка получается в нижней части



2.1 Предварительная обработка

Каждый кадр эталонной и обработанной видеопоследовательности пространственно фильтруется фильтром нижних частот и подвергается субдискретизации до 3 различных разрешений, R1, R2, R3:

	оригинальный кадр		R1		R2		R3
высота x ширина	1 080 × 1 920	→	540 × 960	→	270 × 480	→	96 × 128

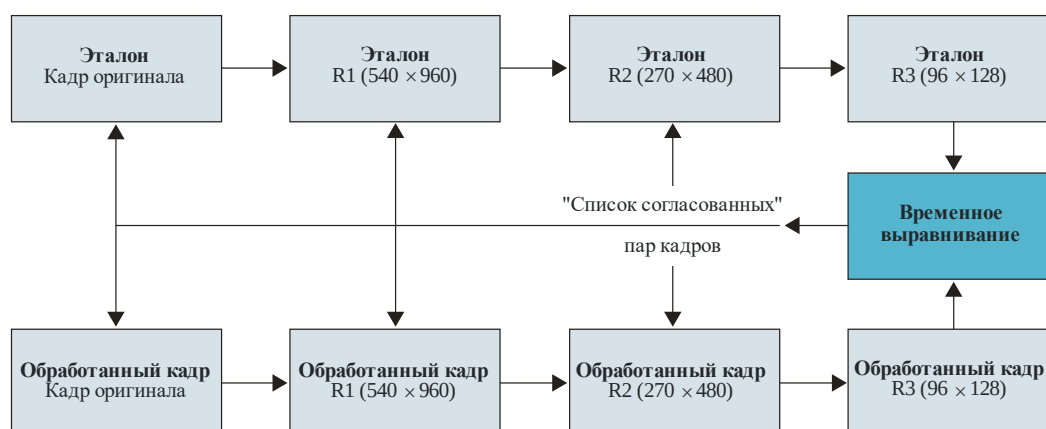
См. метод `CFrameAnalysisFullRef::ContentTimeAlignment`, генерирующий кадры с разрешением R3 и `CFrameSeq::ReadFrame` для генерации кадров с разрешением R1 и R2.

Следует отметить, что реализация не совсем проста из-за ограничений по памяти.

На рисунке 4 показана схема с использованием трех значений разрешения, полученных с помощью субдискретизации.

РИСУНОК 4

Кадры эталонной и обработанной видеопоследовательности фильтруются фильтром нижних частот и подвергаются субдискретизации до 3 различных разрешений. Наименьшее разрешение R3 используется для выполнения временного выравнивания кадра. Результирующий список выровненных кадров может быть использован для приведения в соответствие кадров для любого другого разрешения



ВТ.1907-04

2.2 Временное выравнивание

Временное выравнивание осуществляется путем использования эталонной и обработанной видеопоследовательностей при низком разрешении R3.

Временное выравнивание выполняется рекурсивным образом, как описано ниже:

- 1) найдите "якорный" кадр в эталонной последовательности (`Ref_anchor`).
- 2) этот кадр согласуйте с наиболее подходящим кадром из искаженной последовательности (`Deg_best_match`).

Возьмите этот наиболее подходящий кадр в искаженной последовательности (`Deg_best_match`) и сопоставьте его с кадрами, близкими к "якорному" кадру эталона (`Ref_anchor`). Попытайтесь найти лучшее соответствие по критерию подобия, между `Deg_best_match` и кадрами в среде `Ref_anchor` и сохранить выбранные кадры как наиболее подходящие пары. В качестве критерия подобия между Y-плоскостью обработанного кадра x и эталонного кадра y используется функция:

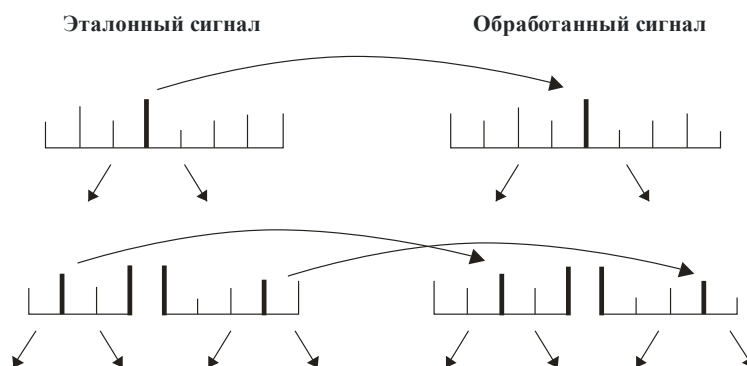
$$\text{sim} = \exp(-\text{mean_square_diff}(a \cdot x + b, y)) \quad (2.1)$$

с параметрами a , b , выбранными так, чтобы среднеквадратичная разность между значениями Y-плоскости обработанного кадра x и эталонного кадра y была минимизирована, см. метод `FrameSimilarity::similarity` в эталонной реализации.

- 1) Если эти подобранные по соответствию пары кадров хорошо согласуются (критерий подобия прошел порог принятия), разделите эталонные и обработанные видеопоследовательности на соответствующие пары кадров, каждая в двух видеопоследовательностях до и после находящихся в соответствии кадров. Начать с п. 1) для обеих пар эталонной и искаженной подпоследовательностей.
- 2) Если соответствующие пары кадров плохо согласуются друг с другом, то начните снова с п. 1) с другим "якорным" кадром эталонной видеопоследовательности. Поскольку нет априорного знания ожидаемого значения "хорошего" подходящего кадра, порог соответствия многократно понижается. Следующие значения были определены на основе многих проверочных выборок данных: начальный порог по отношению к критерию подобия уравнения (2.1) составляет 0,98. После неудачной попытки согласования с 10 якорными кадрами он понижается на коэффициент 0,98, и процесс согласования перезапускается с п. 1). Таким образом, делается до 10 дальнейших попыток выбора якорного кадра, и если они не подходят, то предел снова снижается. Это продолжается до достижения минимального значения 0,1. См. `SQ_TimeAlignment::findAnchorAndDescend` для получения подробных данных о реализации.

РИСУНОК 5

Иллюстрация рекурсивного подхода, используемого для временного выравнивания. Якорный кадр эталонной последовательности согласуется с кадром обработанной последовательности. Затем обе последовательности разделяются и в каждой подпоследовательности выбирается и согласуется якорный кадр



ВТ.1907-05

Результатом временного выравнивания является последовательность (по существу, "список соответствия"), присваивающая каждому кадру обработанной видеопоследовательности кадр эталонной последовательности, или указатель, означающий, что достаточно хороший соответствующий кадр не может быть найден. Таким образом, для более поздних стадий обработки каждый согласованный кадр обработанной видеопоследовательности имеет соответствующий кадр эталонной последовательности. Те кадры обработанной видеопоследовательности, которые имеют указатель "не согласован", будут сравниваться с двумя кадрами эталонной последовательности, соответствующими предыдущему и последующему "согласованным" кадрам обработанной видеопоследовательности. Следует отметить, что "предел соответствия" выбирается весьма низким, так что только очень сильно искаженные кадры имеют указатель "не согласован".

См. метод `CFrameAnalysisFullRef::sqVTA_ContentFrameTimeAlignment_M` для получения всех подробных данных о реализации.

2.3 Пространственное выравнивание кадра

Выполните итерацию всех кадров обработанной видеопоследовательности и:

- 1) Если данный кадр не согласован, то используйте предыдущее пространственное выравнивание. Если данный кадр согласован, выполните пространственное выравнивание между обработанным и соответствующим ему – исходя из списка согласований временного выравнивания – эталонным кадром:

- а) для первого кадра, инициализируйте пространственный сдвиг, который должен быть равным 0 (в горизонтальном и вертикальном направлениях). Для последующих кадров, используйте в качестве предварительного, пространственное выравнивание предыдущих согласованных кадров.
- б) выполните итерацию всех возможных пространственных сдвигов (горизонтальных и вертикальных), используя предел, указанный в пункте 2) ниже. Если другой пространственный сдвиг приводит к значительно (по отношению к функции стоимости) меньшей разнице между обработанным и соответствующим эталонным кадром, пространственный сдвиг корректируется. В качестве функции стоимости используется функция:

$$\text{rmse}(Y(\text{dv}, \text{dh}), Y_{\text{ref}}) + \text{abs}(\text{dv}) + \text{abs}(\text{dh}),$$

где Y обозначает Y -плоскость обработанного кадра с разрешением $R1$ и Y_{ref} обозначает эталонный кадр с разрешением $R1$, $Y(\text{dv}, \text{dh})$ обозначает кадр Y , сдвинутый на dv и dh , где dv , dh – вертикальные и горизонтальные сдвиги. Второй и третий члены, включены в функцию стоимости для учета малых пространственных сдвигов. Следует отметить, что небольшая граница кадров пропускается для расчета rmse , чтобы избежать более сложной обработки границ.

- в) Таким образом, изменяющиеся во времени пространственные сдвиги могут быть компенсированы. Ошибочное выравнивание в одном кадре может быть исправлено путем выравнивания последующих кадров.
- 2) Этот первый этап автоматизированного выравнивания пространственного сдвига ограничен ± 4 пикселями. Что касается более значительных пространственных сдвигов, см. п. 2.9.
 - 3) После выполнения пространственного выравнивания каждый кадр в обработанной видеопоследовательности имеет соответствующий кадр эталонной последовательности (или два кадра в случае несогласованного кадра) в соответствии со списком согласований при временном выравнивании и четкую коррекцию пространственного сдвига. Таким образом, кадры обработанной видеопоследовательности могут точно сравниваться с эталонными кадрами. Это является основополагающим аспектом для последующего выделения характерных признаков.

См. метод `CFrameAnalysisFullRef::DetermineSpatialAlignment` для получения всех подробных данных о реализации. Постоянный порог на шаге 2 (± 4 пикселя) может быть увеличен для обеспечения более значительных пространственных сдвигов.

2.4 Определение признаков локального сходства и локального различия

Для каждой выравниваемой пары кадров определяется набор пространственных характеристик качества.

Сначала вычисляется мера локального *сходства* и *различия* путем итерации в примыкающих, равномерно распределенных квадратных областях размером 13×13 обработанного и эталонного кадров при разрешении $R2$. Так как размер разрешения $R2$ не делится на 13, небольшая граница игнорируется.

Локальные области называются s_{prc} и s_{ref} , а сходство S и различие D вычисляются по формулам:

$$S = (\text{cor}(s_{\text{prc}}, s_{\text{ref}}) + 25) / (\text{var}(s_{\text{ref}}) + 25) \quad (4.1)$$

$$D = \sqrt{\text{avg}((S * (s_{\text{prc}} - \text{mean}(s_{\text{prc}})) - (s_{\text{ref}} - \text{mean}(s_{\text{ref}})))^2))}, \quad (4.2)$$

где cor – это корреляция и var – это дисперсия количества пикселей в соответствующей квадратной области. Функция avg вычисляет среднее по всем пикселям квадратной области, а $\sqrt{}$ обозначает квадратный корень. Значения D и S являются основным источником формирования пространственных характеристик качества.

На данном этапе, характеристики сходства и различия S , D являются матрицей значений для каждого кадра, причем одно значение соответствует каждому квадрату местной области. Важным показателем для воспринимаемого качества является не только среднее значение, но форма распределения значений S , D соответственно.

2.5 Анализ распределения локальных признаков

Этот раздел начинается с введения некоторых обозначений:

Пусть *квантиль* (X, c) обозначает *c-квантиль* распределения значений (входов) вектора или матрицы X . Точнее, для вектора X и константы c при $0 \leq c \leq 1$, квантиль:

$$q = \text{quantile}(X, c)$$

является значением q , при котором дробная часть c всех значений X меньше или равна q .

Функция *trimmed_mean* определяется следующим образом. Это обозначение будет использоваться в дальнейшем. Для матрицы x , усеченное среднее:

$$\text{trimmedMean}(X, c)$$

является средним всех входных значений x между c и $(1-c)$ квантилями x .

Например, $\text{trimmedMean}(X, 0.1)$ – это среднее всех значений X , игнорируя 10% наименьших и 10% наибольших значений x .

Запись $X(X > c)$ обозначает совокупность всех значений X , больших чем c . Например:

$$\text{trimmedMean}(X, c) = \text{mean}(X(X > \text{quantile}(X, c) \text{ и } X < \text{quantile}(X, 1-c))).$$

Используя эти обозначения, вычисляются следующие значения характеристик на основе значений s из уравнения (4.1) и D из уравнения (4.2):

$$s_m = \text{trimmedMean}(S, c) \quad (5.1)$$

$$d_m = \text{trimmedMean}(D, c) \quad (5.2)$$

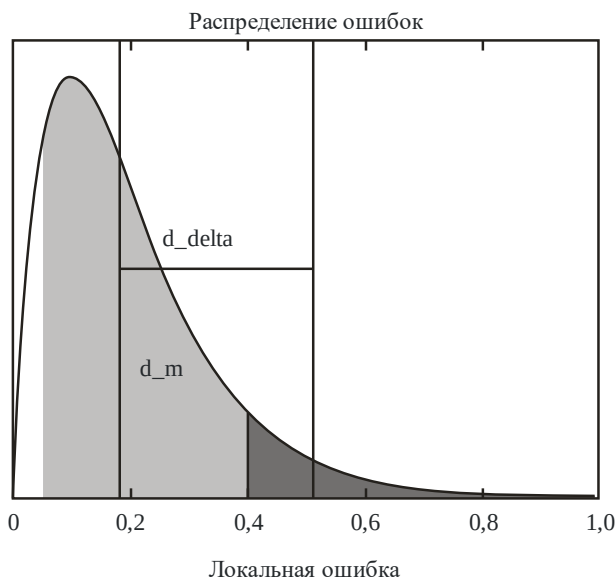
$$s_delta = s_m - \text{mean}(S(S < \text{quantile}(S, c))) \quad (5.3)$$

$$d_delta = \text{mean}(D(D > \text{quantile}(D, 1-c))) - d_m, \quad (5.4)$$

используя $c = 0,2$. Это визуально показано на рис. 6, в котором представлены значения d_m и d_delta .

РИСУНОК 6

Показано распределение локальной характеристики D . Усеченное среднее d_m соответствует среднему области светло-серого цвета (черная вертикальная линия).
Значение d_delta соответствует разнице средних значений в темно-серой и светло-серой областях (горизонтальная линия)



ВТ.1907-06

См. метод `CFrameAnalysisFullRef::ComputeSimilarity` для вычисления значений S и D .

2.6 Вычисление характеристики блочности

Характеристика *блочности* вычисляется с использованием кадров при разрешении R1. Эта характеристика измеряет видимость границ блоков, появившихся в результате кодирования и/или ошибок при передаче. В связи с вычислением при разрешении R1, автоматически устанавливается акцент на восприятие видимости краев. Начиная с обзора, функция блочности вычисляет:

- 1) Производные по направлению (по краю изображения) для горизонтального и вертикального краев. В результате получаются две матрицы (одна для горизонтального и одна для вертикального края) для каждого кадра видеопоследовательности, называемые `verGrad_n` и `horGrad_n` в приведенном ниже псевдокоде.
- 2) сумму логарифмов по строкам и колонкам горизонтального и вертикального краев, при этом в результате получаются два вектора, один соответствует сумме горизонтальных краев, а другой соответствует сумме вертикальных краев, обозначаемых как `sumW` и `sumH` ниже.
- 3) среднее значение подвыборки `sumW` и `sumH`, соответственно, на размере шага `n` и смещении `m`, вычисленное с помощью функции `vq_AvgSubsample` описанной ниже.

Идея состоит в том, что устойчивая структура блоков размером `n` показывает, как важна разница `delta_edge` в `vq_AvgSubsample` на размере шага `n`, вычисленная для различных смещений.

Например блочная структура размером 4 в кадре оригинала имеет блочную структуру размером 2 при разрешении R1. Таким образом, вычисление `vq_AvgSubsample(x, 2, 0)` и `vq_AvgSubsample(x, 2, 1)` должно показать существенное различие, если имеет место устойчивая блоковая структура. Чтобы избежать зависимости от контента, эксперименты с использованием большой выборки видеопоследовательностей показали, как соотносить вычисленную разницу, измеренную для обработанной видеопоследовательностей, со значениями эталонной последовательности.

Более подробную информацию о расчете лучше всего объяснить с помощью следующего псевдокода. Здесь `horGrad` и `verGrad` являются горизонтальной и вертикальной пространственными производными кадра, заданными как разница между соседними пикселями,

```
verGrad_n(i,j) = Y_n(i+1,j) - Y_n(i,j); и
horGrad_n(i,j) = Y_n(i,j+1) - Y_n(i,j),
```

где `Y_n(i, j)` обозначает значение пикселя в позиции `(i, j)` в `Y`-плоскости кадра `n`. Функция

```
vq_AvgSubsample( x, step, offset )
```

вычисляет среднее значение вектора `x` на всех выборках с размером шага `step` и начиная со смещения `offset`.

```
// образуйте замкнутый контур по всем кадрам и вычислите:
for( UINT i=0; i<horGrad.Height; i++){
    for( UINT j=0; j<horGrad.Width; j++){
        w = (double)verGrad(i,j);
        h = (double)horGrad(i,j);
        // сумма краев (-2: небольшие различия могут быть результатом целых
        // значений, используемых для сохранения кадров)
        sumW(i) += log(1.0+max(0.0,fabs(w)-2.0));
        sumH(j) += log(1.0+max(0.0,fabs(h)-2.0));
    }
}

double dH0 = vq_AvgSubsample( sumH, 2, 0 );
double dH1 = vq_AvgSubsample( sumH, 2, 1 );
double dW0 = vq_AvgSubsample( sumW, 2, 0 );
double dW1 = vq_AvgSubsample( sumW, 2, 1 );

edge_max = 0.5 * (vq_Max(dW0,dW1) + vq_Max(dH0,dH1) );
edge_min = 0.5 * (vq_Min(dW0,dW1) + vq_Min(dH0,dH1) );
```

```

// теперь: обозначьте через edge_max(i) вышеуказанное значение of edge_max,
соответствующее

// кадру i обработанной видеопоследовательности и через edge_max_ref(i) вышеуказанные
значения

// edge_max, соответствующие кадру i эталонной видеопоследовательности,
// и аналогичным образом для edge_min(i), edge_min_ref(i). Затем вычислите:

for( UINT i=0; i<nbOfFramesInProcessedVideo; i++ ){
    // определите кадр nb эталонного кадра (в соответствии со списком согласований)
    UINT i_ref = (UINT)floor(ref_frameNb_all(i)+0.5f);

    float delta_edge = edge_max(i) - edge_min(i);
    float delta_edge_ref = edge_max_ref(i_ref) - edge_min_ref(i_ref);

    x(i) = vq_Max(0.0f, delta_edge - delta_edge_ref)/(1.0f+edge_max(i));
}
// блочность(i) является в таком случае нелинейным монотонным преобразованием x(i) ...

```

Следует отметить, что в связи с возможностью увеличения квантования до 720-кадров, расчет несколько усложняется:

См. `vquad_hd::vq_BlockinessPhaseDiff` и `CQualityModelFullRef::Blockiness` для получения всех подробных данных о вычислении характеристик блочности.

2.7 Вычисление характеристики искажений типа "движение толчками" (временное качество)

Характеристика искажений типа "движение толчками" вычисляется путем усреднения произведения относительного времени отображения, нелинейного преобразования времени отображения и нелинейного преобразования интенсивности движения. Интенсивность движения в основном появляется из-за межкадрового рассогласования по отдельным областям кадра. *Время отображения* это – время в миллисекундах, в течение которого кадр отображается на экране. Следует отметить, что с целью определения времени отображения каждого кадра, осуществляется локальный анализ интенсивности движения, поскольку кадры в обработанной видеопоследовательности могут быть повторениями предыдущих кадров.

Характеристика движений толчками учитывает количество пропущенной информации во время воспроизведения обработанной видеопоследовательности. Этот показатель весьма небольшой в случае свободно воспроизводимой последовательности, в то время как он увеличивается в случае паузы или пониженной частоты кадров. С другой стороны, для фиксированного временного ухудшения характеристика движения толчками имеет более важное значение для видеопоследовательностей с большей интенсивностью движения.

Следующий псевдокод показывает подробную информацию. Следует отметить, что входными параметрами являются вектор интенсивности движения `motionInt`, вектор вероятностей повторения кадров `repFrame` и вектор времени отображения кадров `displayTime`. В результате на выходе появляется вектор `jerkiness` характеристики движения толчками для каждого кадра обработанной видеопоследовательности. Более подробно, вектор `motionInt` означает среднеквадратичную межкадровую разность, измеренную на Y плоскости при разрешении R2. Вектор `repFrame` означает вероятность повторения кадров, т. е. в зависимости от интенсивности движения, каждый кадр имеет вероятность быть повторением предыдущего кадра: в случае идеального повторения предыдущего кадра, реальный кадр имеет вероятность 1 для события повторения этого кадра. В случае большой интенсивности движения реальный кадр имеет вероятность 0 того, что будет иметь место повторение предыдущего кадра. Промежуточные значения вероятностей могут иметь место, если интенсивность движения имеет небольшие, но ненулевые значения. Имеем,

$$\text{repFrame}(i) = \begin{cases} 0 & \text{if } m(i) < p/2 \\ (m(i) - p/2) / p & \text{if } p/2 \leq m(i) < 3/2 * p \\ 1 & \text{if } 3/2 * p \leq m(i), \end{cases}$$

где $m(i)$ обозначает степень интенсивности движения кадра i . Эмпирически было выбрано значение параметра $p = 0,01$.

```
int vq_CalcJerkiness( const CVector<float> & motionInt,
                    const CVector<float> & repFrame,
                    const CVector<float> & displayTime,
                    CVector<float> & jerkiness ){

    // -----
    // 4 параметра показателя "движение толчками": определяемые с использованием
    // большой выборки видеопоследовательностей, содержащих только
    // временные искажения.
    float a = 0.9f;
    float b = 5.0f;

    float aT = 40.0f;
    float bT = 5.0f;
    // -----

    // определите вероятность нового кадра = 1 - вероятность повторенного кадра:
    CVector<float> newFrame = repFrame*(-1.0f) + 1.0f;

    // подсчитайте количество неповторенных кадров
    float fNbRepeated = repFrame.Sum();
    float fNbNonRepeated = repFrame.Length() - fNbRepeated;

    // вычислите показатель "движение толчками"
    float fR = 0.0f;
    // определите интервалы между кадрами(≈ время отображения)
    // длиной i
    for( UINT i=1; i<= iNbFrames; i++ )
    {

        // определите период повторения кадров, начиная с положения j
        for( UINT j=0; j<iNbFrames-i+1; j++ )
        {
            float fP = newFrame(j); // prob. : start of repetition block
            for( UINT k=1; k<i; k++ )
            {
                fP *= repFrame(j+k); // prob. : all repeated frames
            }
            if( i+j < iNbFrames )
            {
                fP *= newFrame(j+i); // prob. : end of repetition block
            }

            // вычислите время отображения(в сек.) кадра j,
            // если он отображается от
            // момента t_j до момента t_(j+i), который наблюдается с вероятностью fP
            float fDispTime = displayTime.SumPart(j,j+i)/1000.0f;

            // -> измерьте значение jerking и добавьте к результату
            float fIFDiff = motionInt( j+i-1 );

            // нормирование значений: так чтобы нулевое значение jerkiness
            соответствовало 0,
            // а насыщение происходило на 1

```

```

float c = 1.0f/(1.0f+exp(b));
float cT = 1.0f/(1.0f+exp(bT));
float fJ = 1.0f/(1.0f + exp( -( a * fIFDiff - b) ));
float fJT = 1.0f/(1.0f + exp( -( aT * fDispTime - bT)));
fJ = (fJ - c)/(1.0f - c);
fJT = (fJT - cT)/(1.0f - cT);

// значение jerkiness: вероятность * interframeDiffFactor *
// displayTimeFactor

fR = fP * fJ * fJT * fDispTime;

// добавьте к вектору jerkiness в положении j+i (соответствующем
// окончанию времени отображения)
jerkiness( vq_Min(j+i,iNbFrames-1) ) += fR;
}
}
return 0;
}

```

См. метод `vquad_hd::vq_ProbOfRepeatedFrame` и `vquad_hd::vq_CalcJerkiness` для получения подробных данных о реализации.

2.8 Объединение с MOS (показателем усредненного мнения)

Характеристики, описанные выше; сходство, определенное с помощью s_m и s_{Δ} ; различие, определенное с помощью d_m и d_{Δ} ; эффект блочности и эффект движения толчками являются основой для оценки результирующего показателя, вместе с вектором времени отображения кадров `displayTime`. Эти векторы имеют длину, равную количеству кадров обработанной видеопоследовательности.

Чтобы отобразить эти характеристики на шкалу восприятия, используется параметризованная S-образная функция:

$$S: \mathbb{R} \rightarrow \mathbb{R}, \quad y = S(x).$$

Функция S определяется тремя параметрами (p_x , p_y , q). Эти параметры имеют следующее толкование: (p_x, p_y) являются расположением в $\mathbb{R} \times \mathbb{R}$ и q – это наклон точки перегиба кривой, более детально:

$$S(x) = \begin{cases} a * x^b & \text{if } x \leq p_x \\ d / (1 + \exp(-c * (x - p_x))) + 1 - d & \text{else,} \end{cases} \quad (8.1)$$

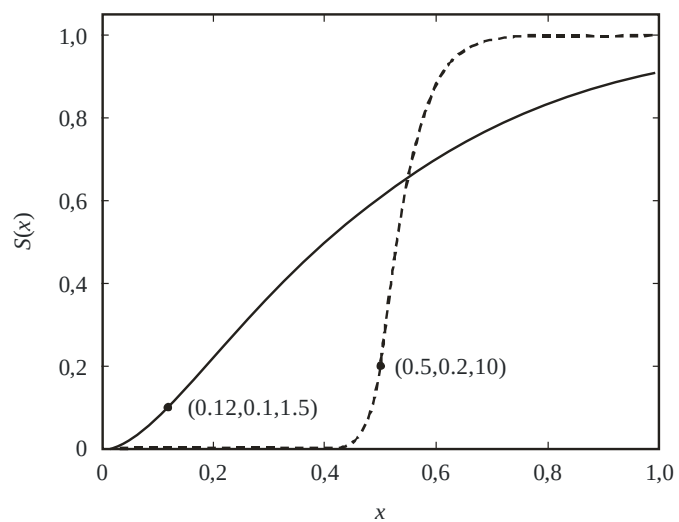
где:

$$\begin{aligned} a &= p_y / p_x^{(q * p_x / p_y)} \\ b &= q * p_x / p_y \\ c &= 4 * q / d \\ d &= 2 * (1 - p_y). \end{aligned}$$

См. рис. 7, где изображен график S-функции с различными параметрами. S-образная функция начинается с нуля, растет полиномиально до точки перегиба и экспоненциально стремится к насыщению на уровне 1.

РИСУНОК 7

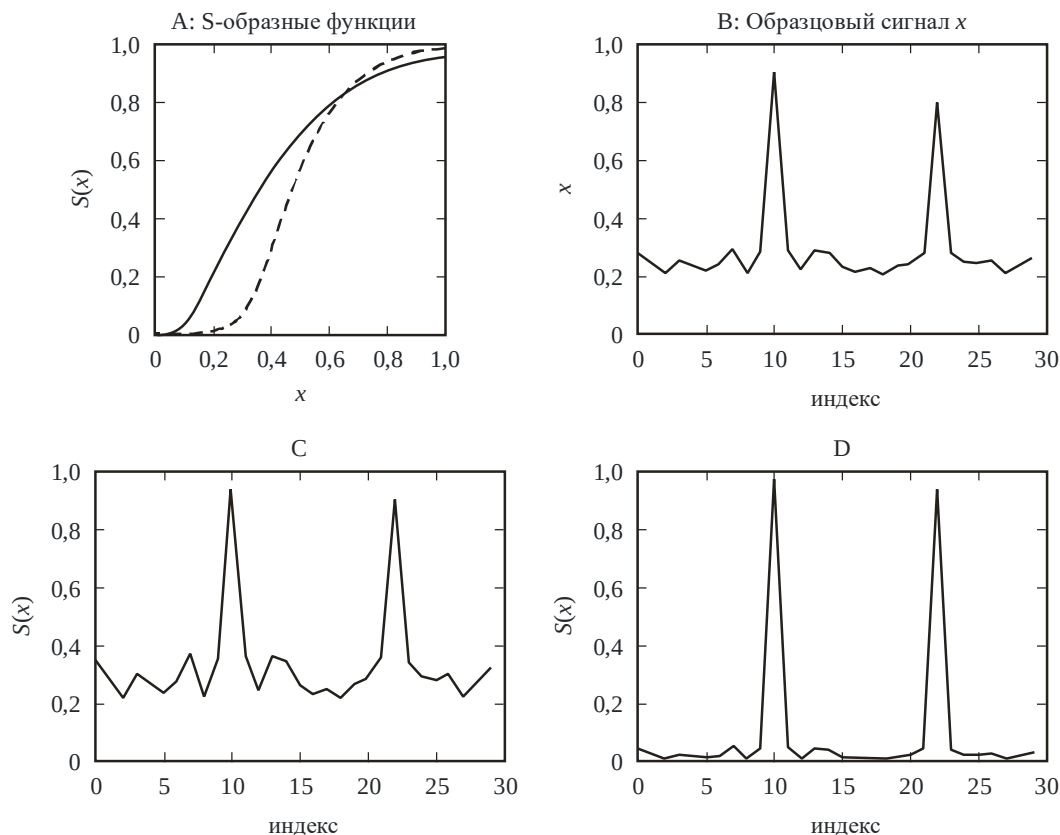
S-образные функции, параметризованные положением и наклоном точки перегиба. Показаны две примерные функции для различных параметров



ВТ.1907-07

РИСУНОК 8

График (А) показывает S-образную функцию для 2 различных наборов параметров. График (В) показывает вектор сигнала выборки. Все значения x преобразуются с помощью двух S-образных функций (А). Результат показан на рис. (С) и (D). S-образное преобразование используется для отображения характеристики x на шкалу восприятия с помощью параметров, подходящих для выборок данных, показанных на (С). Используя S-образную функцию (пунктирная линия на (А)) с параметрами, зависящими от данных (первый параметр устанавливается на несколько средних значений сигнала, x), может быть построен детектор переходных искажений (D)



ВТ.1907-08

Значения характеристики могут быть отображены на шкалу восприятия при использовании S-образной функции с постоянными параметрами. Посредством параметров, зависящих от данных, S-образная функция может использоваться, например, для сжатия всех значений в диапазоне от ниже среднего до малых значений, и расширения всех более высоких значений. Таким образом, она может быть использована для вычисления характеристики, чувствительной к переходным искажениям, см. рис. 8.

Чтобы продолжить описание модели, во-первых, определите степень ухудшения на основе определенных выше признаков сходства, но придавая больший весовой коэффициент сильным ухудшениям:

$$d_s = 1 - s_m + 1.5 s_{\text{delta}}.$$

Следующая идея заключается в использовании двух S-образных функций, первая, использующая фиксированный набор параметров cod_par , для преобразования d_s в d_{cod} на шкале восприятия, связанной с основным искажением, отражающим ошибки из-за видеокодирования.

Вторая S-образная функция с параметрами, зависящими от данных, служит для преобразования d_s в d_{trans} на шкале восприятия, связанной с переходными искажениями, отражающими ошибки передачи.

Более детально, относительно указанного ниже псевдокода, – вызовите функцию `SplitCodTrans`, используя в качестве входного параметра вектор d_s и вектор времени отображения кадра `disp_time` и получите на выходе d_{cod} , d_{trans} ,


```
SplitCodTrans(d_s, disp_time, d_cod, d_trans).
```

Следующий псевдокод показывает подробные данные о функции. Следует отметить, что

```
stat.STransform(x, px, py, q)
```

обозначает S-образную функцию и имеет в качестве входных данных реальное значение x , которое будет преобразовано, и три параметра, обозначаемые как (px, py, q) в уравнении (8.1).

```
SplitCodTrans( const CVec& v, const CVec& dispTime,
               CVec& v_cod, CVec& v_trans ){

    // эти параметры определяются эмпирическим путем
    float qPosSmall = 0.55f;
    float qPosLarge = 0.65f;

    // q это среднее для значений в v между квантилями qPosSmall
    // и qPosLarge
    float q = r.TrimmedMean( qPosSmall, qPosLarge, v, dispTime );

    for( UINT i=0; i<v.Length(); i++ ){

        // используемые здесь параметры являются результатом
        // согласования с выборочными данными
        v_cod(i) = stat.STransform(v(i), 0.07f, 0.1f, 2.0f);

        // Следует отметить, что STransform напрямую применяется не к v, а
        // к v(i)-q . Такому подходу здесь отдается предпочтение с учетом
        // числовых соображений результирующего значения STransform.

        // является частью вышеуказанного параметра v значения квантиля q
        v_trans(i) = vq_Max(0.0f, v(i)-q);

        float px = 0.5f * (q+0.2f);
        // используемые здесь параметры являются результатом
        // согласования с выборочными данными
        v_trans(i) = stat.STransform(v_trans(i), px, 0.1f, 16.0f);
    }
}
```

См. CQualityModelFullRef::SplitCodTrans для получения всех подробных данных о реализации.

По аналогии, значения d_diff_cod , d_diff_trans выводятся из d_m , d_delta , путем установки

```
d_diff = d_m + 1.5 d_delta,
```

и вызова функции

```
SplitCodTrans (d_diff, disp_time, d_diff_cod, d_diff_trans),
```

используя три параметра для двух S-преобразований

```
cod_par = (4.0f, 0.05f, 0.2f)
```

```
trans_par = (0.5*(q+4.0f), 0.1f, 0.4f),
```

где q обозначает среднее между квантилями, как и в приведенном выше псевдокоде.

Выходными значениями функции являются d_diff_cod , d_diff_trans .

Далее, вычисляется значение характеристики, связанное с переходными большими значениями эффекта движения толчками, используя S-образное преобразование этого эффекта:

```
d_t_trans = S(max(0, jerkiness-q))
```

с параметрами S-образного преобразования, заданными как $(\max(0.048, q), 0.2, 40.0)$, где q обозначает межквантильное среднее между 0,55 и 0,65 квантилями вектора эффекта движения толчками. Параметры определялись путем согласования с большим набором данных выборки.

См. `CQualityModelFullRef::SplitTempTrans` для получения полной информации о реализации.

Далее, базовое качество q_{cod} определяется как:

$$q_{\text{cod}} = (1 - d_{\text{cod}}) * (1 - d_{\text{diff_cod}}) * (1 - \text{blockiness}) \quad (8.2)$$

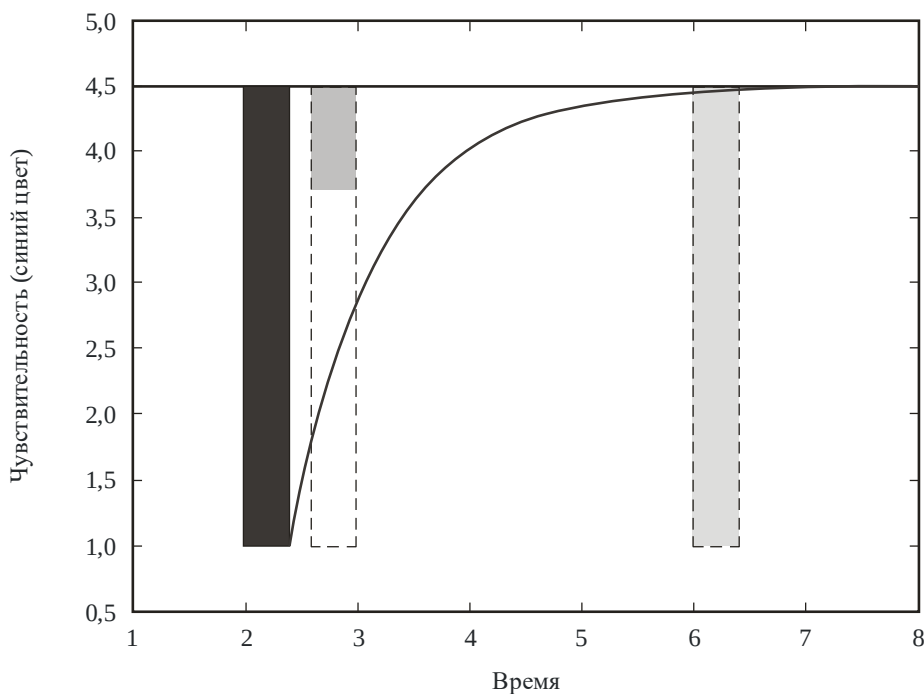
а переходное качество определяется как:

$$q_{\text{trans}} = (1 - d_{\text{trans}}) * (1 - d_{\text{diff_trans}}) * (1 - d_{\text{t_trans}}).$$

Влияние дополнительного ухудшения уменьшается, если это происходит вскоре после первого ухудшения. Для учета этого эффекта, q_{trans} трансформируется в q_{fq} . Данная идея иллюстрируется на рис. 9.

РИСУНОК 9

Сплошная черная полоса указывает на ухудшение, имеющее место в гипотетической видеопоследовательности. Черная сплошная линия указывает на чувствительность к последующим ухудшениям. Влияние дополнительного ухудшения в степени, отмеченной пунктирной линией, снижается (первая темно-серая полоса), если это происходит вскоре после первого ухудшения. Влияние последующего ухудшения не изменится, если оно происходит с большим временным интервалом между ними (правая светло-серая полоса)



ВТ.1907-09

Подробные данные о вычислениях лучше всего описываются следующей секцией псевдокода, которая использует $1 - q_{\text{trans}}$ в качестве входного вектора v , вектор `disp_time` является вектором времени отображения кадра. Постоянная времени спада $dT=1000$ мс была определена с помощью выборочных данных, затем положим

$$q_{\text{fq}} = 1 - w, \quad (8.3)$$

где w является выходным вектором функции:

```

DegFreq( const CVec& v, const CVec& disp_time, CVec& w, float dT ){
    // постоянная времени для временной интеграции искажений
    float t_const = 80.0f;
    w(0) = v(0)*vq_Min(t_const,disp_time(0))/t_const;

    for( UINT i=1; i<v.Length(); i++ ){
        // объедините искажения за последние t_const миллисекунд:
        // используйте функцию указателя, который равен 1 в интервале
        // [t_i-t_const, t_i] и 0 в других случаях
        float dT_sum = 0.0f;
        UINT j=0;
        float v_sum = 0.0f;
        while( dT_sum < t_const && (int)i-(int)j>=0 ){
            // b частично совпадает в интервале [0,1] времени отображения
            // кадра i-j по отношению к интервалу интеграции
            //      t_i-t_const t_i
            //      |         |
            //      |         |-----> окно интеграции
            // -----> время
            // | | | | | | | кадры
            //
            // ->| b*dT |<-
            //
            float b = vq_Min(t_const-dT_sum,disp_time(i-j))/t_const;

            v_sum += v(i-j)*b;
            dT_sum += disp_time(i-j);
            j++;
        }
        // вычислите коэффициент спада a:
        float a = exp(-disp_time(i-1)/dT);
        // искажение в этом случае включает:
        // 1) линейную комбинацию спада предыдущего искажения
        // и текущего(суммарного) искажения, или
        // 2) суммарное текущее искажение (если оно сильнее, чем вариант 1) ).

        w(i) = vq_Max(v_sum, a * w(i-1) + (1.0f-a) * v_sum);
    }
}

```

Процесс усреднения

Пусть $\text{disp_time}(i)$ обозначает время отображения кадра i и $\text{jerkiness}(i)$ значение эффекта движения толчками, соответствующего кадру i . Выберем q_{cod} из уравнения (8.2) и q_{fq} из уравнения (8.3). Положим

$$\begin{aligned}
 Q_t &= 1 - 1/T \sum_i \text{jerkiness}(i) \\
 Q_{\text{fq}} &= 1/T \sum_i q_{\text{fq}}(i) * \text{disp_time}(i) \\
 Q_{\text{cod}} &= 1/T \sum_i q_{\text{cod}}(i) * \text{disp_time}(i),
 \end{aligned}$$

где $T = \sum_i \text{disp_time}(i)$, и $\sum_i$ обозначает сумму по всем индексам $i=0, \dots, \text{number_of_frames}$.

Окончательная прогнозируемая оценка является произведением и масштабируется в соответствии с диапазоном MOS [1,5]:

$$s = 4 * Q_t * Q_{\text{cod}} * Q_{\text{fq}} + 1.$$

См. метод `CQualityModelFullRef::PredictScoreCodTrans` для получения более подробных данных относительно прогнозирования оценки.

2.9 Обработка видеопоследовательностей со значительным рассогласованием в пространственном выравнивании

Чтобы избежать неправильного прогнозирования в случае относительно большого пространственного рассогласования между эталонной и обработанной видеопоследовательностями, вышеупомянутые шаги рассчитываются для трех различных горизонтальных и вертикальных пространственных шагов выравнивания видеопоследовательности, и максимальная прогнозируемая оценка по всем пространственным положениям берется как показатель окончательной оценки качества.

В каждом направлении используется размер шага из четырех пикселей. Таким образом, реализуется пространственный диапазон поиска в ± 8 пикселей. Это значение легко покрывает максимальный используемый пространственный сдвиг в наборе тестов (± 5 пикселей). Поскольку это расширение осуществляется на функции высокого уровня при моделировании, диапазон выравнивания может быть легко адаптирован к любым более значительным сдвигам или может быть уменьшен (например, ± 4 пикселей) для экономии вычислительных ресурсов.

См. `vquad_hd :: vq_vquad08`.

2.10 Реализация эталонного исходного кода

В этом разделе содержится встроенный архив с исходным кодом на C++, охватывающим основные части и функции для соответствующего описания реализации данной модели. Все указания на конкретные реализации, сделанные в предыдущих разделах, ссылаются на этот эталонный исходный код.

Приложение 3

Проверка на совместимость

В настоящее Приложение включено цифровое дополнение, содержащее следующую информацию и файлы:

- 1) Шестнадцать коротких видеопоследовательностей с высокой четкостью (HD) (эталонные и искаженные последовательности для восьми тестовых вариантов). Эти последовательности охватывают различные виды искажений и содержания. Они предназначены для тестирования совместимой реализации данной модели, выполненной пользователем настоящей Рекомендации с целью сравнения с эталонной реализацией модели. Видеопоследовательности состоят из нескольких кадров исключительно для снижения емкости памяти. Они не предназначены для любых визуальных тестов, а только для испытаний на совместимость с реализацией модели высокой четкости.
- 2) Прогнозируемые оценки MOS для последовательностей высокой четкости, упомянутые в п. 1) как "HD_ConformanceReferenceResults.xls". Эти оценки были получены при помощи эталонной реализации модели высокой четкости.
- 3) Прогнозируемые оценки MOS для пяти ВЧ баз данных HD общего пользования доступны через VQEG. Такие базы данных могут быть использованы в качестве расширенного теста на совместимость с реализациями этой модели.

Критерии теста на совместимость

- i) Восемь эталонных оценок, приведенные в п. 2), должны точно воспроизводиться при реализации модели. Точное воспроизведение оценки доступно с разрешением в три десятичных цифры.
- ii) Прогнозируемые оценки MOS из пяти баз данных VQEG общего пользования должны воспроизводиться с очень ограниченным отклонением. Незначительные изменения допустимы, так как опыт показал, что различные оптимизации в плане скорости и используемой памяти могут привести к небольшим и незначительным отклонениям в конечной оценке.

ТАБЛИЦА 2

Допустимое распределение различий по всем согласованным тестовым данным

Абсолютная разница	Допустимая доля появлений
> 0,0001	5,00%
> 0,001	1,00%
> 0,01	0,50%
> 0,1	0,05%
> 0,3	0,00%

Для остальных баз данных, отличных от тех, которые определены в настоящем Приложении 3, такое же распределение ошибок не должно превышать. В отношении неизвестных данных для этой статистики должен быть взят набор тестов, по крайней мере из 500 пар файлов – предпочтительно от законченных субъективных экспериментов.

Библиография

- VQEG Final Report of HDTV Phase I Validation Test (2010), “*Video Quality Experts Group: report on the validation of video quality models for high definition video content*”, Video Quality Experts Group (VQEG), <http://www.its.bldrdoc.gov/vqeg/projects/hdtv>
- Рекомендация МСЭ-R ВТ.601-7 (2011 г.), Студийные параметры кодирования цифрового телевидения для стандартного 4:3 и широкоэкрannого 16:9 форматов.
- Recommendation ITU-T J.244 (2008), Calibration methods for constant misalignment of spatial and temporal domains with constant gain and offset.
- Recommendation ITU-R BT.500-12 (2009), Methodology for the subjective assessment of the quality of television picture.
- Recommendation ITU-T J.149 (2004), Method for specifying accuracy and cross-calibration of Video Quality Metrics (VQM).
- Recommendation ITU-T J.247 (2008), Objective perceptual multimedia video quality measurement in the presence of a full reference.
- Recommendation ITU-T J.144 (2004), Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference.
- Recommendation ITU-T P.931 (1998), Multimedia communications delay, synchronization and frame rate measurement.
- Recommendation ITU-T J.148 (2003), Requirements for an objective perceptual multimedia quality model.
- Recommendation ITU-T H.264 (2012), Advanced video coding for generic audiovisual services.
- Recommendation ITU-T J.340 (2010), Reference algorithm for computing peak signal to noise ratio (PSNR) of a processed video sequence with constant spatial shifts and a constant delay.
- Recommendation ITU-T P.910 (2008), Subjective video quality assessment methods or multimedia applications.
- Recommendation ITU-T P.911 (1998), Subjective audiovisual quality assessment methods for multimedia applications.

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