



СУБЪЕКТИВНЫЕ МЕТОДЫ ОЦЕНКИ QoE



КАЧЕСТВО ВОСПРИЯТИЯ, QOE

Субъективные показатели:

- 1. Компоненты восприятия человеком** - культурный фон, мотивация, эмоциональное состояние, внимание и т.д.
- 2. Функции управления сервисом** - опыт работы с конкретной системой и ее уровнем качества, удобство навигации при выборе услуг IPTV, при поиске контента, интуитивно понятный интерфейс.
- 3. Тарифы.**

Объективные показатели:

- 1. Факторы передачи информации** - минимальная скорость передачи данных, максимальный уровень потерь пакетов, задержки и др.
- 2. Факторы функционирования приложений** - параметры кодека, разрешение видеоданных источника, скорость кодирования, схему маскировки ошибок и т.д.
- 3. Факторы услуг** или уровень предоставления услуги - Zapping time, возможность выбора контента, EPG.

СУБЪЕКТИВНАЯ ОЦЕНКА

Этапы тестирования:

1. Выбор или сочетание способов демонстрации видеопоследовательностей;
2. Определение методики сбора мнений экспертов;
3. Выбор методики обработки результатов.

СУБЪЕКТИВНЫЕ МЕТОДЫ ВТ.500-Із

- Методы «***double stimulus***» - последовательное воспроизведение каждой пары видео, состоящей из опорного и тестируемого видео;
- Методы «***single stimulus***» - воспроизведение и оценка только тестируемой последовательности;
- Методы «***stimulus comparison***» - воспроизведение пары видеопоследовательностей и оценка их качества относительно друг друга.

СУБЪЕКТИВНЫЕ МЕТОДЫ. КЛАССЫ

Оценки качества

- Методы, которые позволяют оценить качество системы в оптимальных условиях (т.е. показатели QoS соответствуют норме и лучше).

Оценки ухудшений

- Методы, которые оценивают способность системы поддерживать качество работы в неоптимальных условиях работы (моделируются различные проблемные ситуации на сети).

ШКАЛЫ ОЦЕНОК ДЛЯ МЕТОДОВ SS

- Методы оценок с использованием определенных категорий;
- Численные методы оценок с использованием категорий;
- Методы оценок без использования категорий;
 - с использованием непрерывной шкалы;
 - с использованием цифровой шкалы

Методы субъективной оценки

Методы с двумя источниками воздействия (DSIS, DSCQS, SDSCE)

Альтернативные

С одним источником воздействия (SS)

С использованием определенных категорий (SSCQE)

Численные методы с использованием категорий (SSNSC)

Без использования категорий

Методы оценки качества

Методы сравнения воздействий

С использованием определенных категорий

Без использования категорий

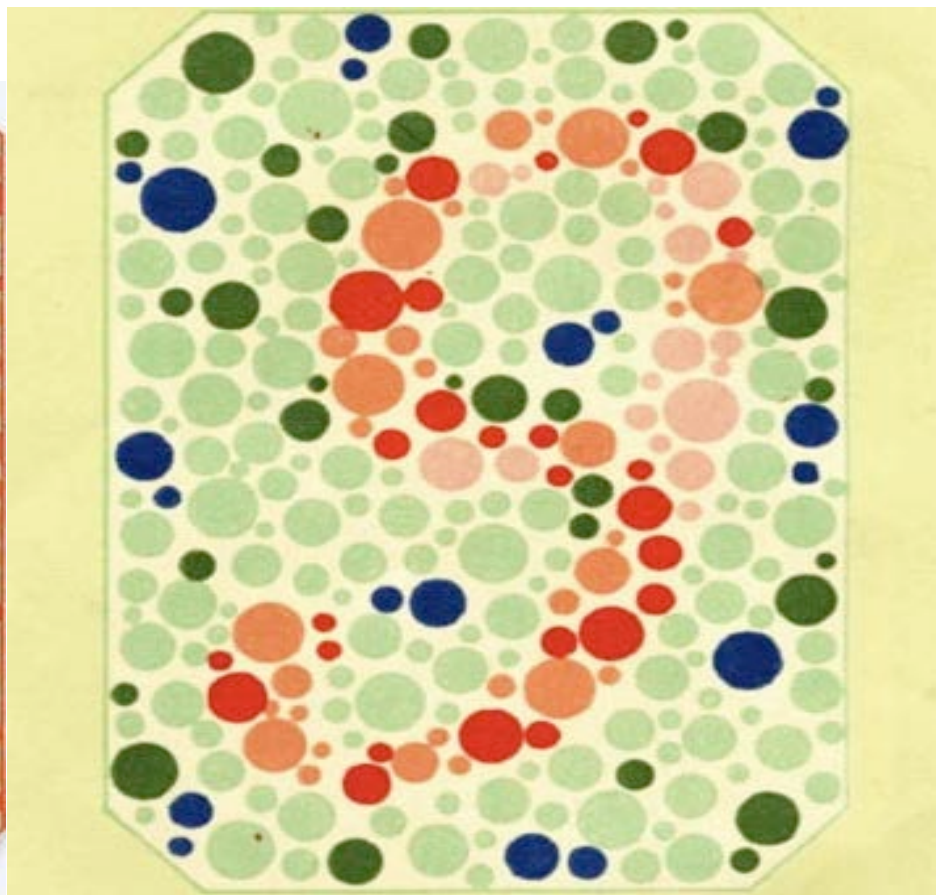
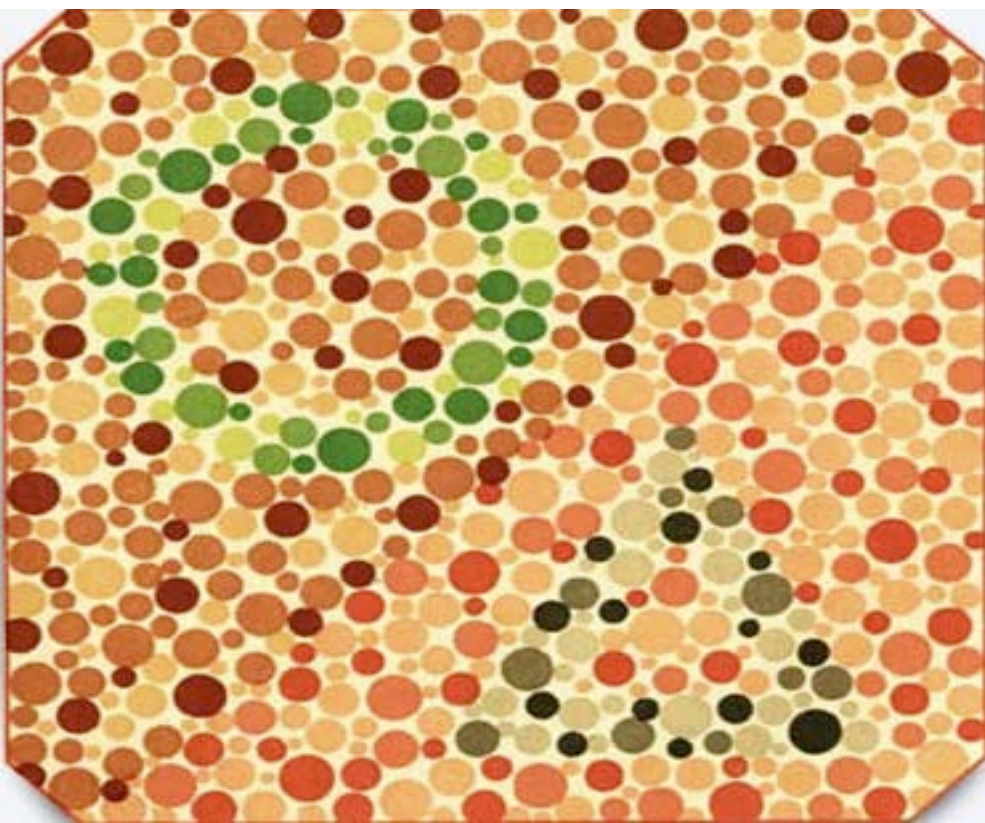
Методы оценки качества

НАБЛЮДАТЕЛИ

- Необходимо, чтобы наблюдателей было не менее 15.
- Они не должны быть экспертами, что означает, что они не связаны непосредственно с оценкой качества телевизионного изображения, как частью их обычной работы, и они не являются опытными оценщиками.
- До сеанса следует оценить (или скорректировать) остроту зрения наблюдателей при помощи
 - таблицы Снеллена и
 - Ландольта
 - и проверить цветовое зрение, используя специально отобранные таблицы (например, Ишихара).
- Необходимое количество наблюдателей зависит от критичности и надежности принятой процедуры испытания и от ожидаемого воздействия оцениваемого эффекта.

Таблицы Ишихара

для проверки цветовосприятия



| | | | |
|-----|--|--|--|
| 0,1 | | | |
| 0,2 | | | |
| 0,3 | | | |
| 0,4 | | | |
| 0,5 | | | |
| 0,6 | | | |
| 0,7 | | | |
| 0,8 | | | |
| 0,9 | | | |
| 1,0 | | | |
| 1,5 | | | |

Таблица Снеллена

Таблица Ландольта

| | | | | | | | | |
|--|--|---------|---------|---------|---------|---------|---------|---------|
| | | VIS=0,1 | | | | | | |
| | | | VIS=0,2 | | | | | |
| | | | | VIS=0,3 | | | | |
| | | | | | VIS=0,4 | | | |
| | | | | | VIS=0,5 | | | |
| | | | | | | VIS=0,6 | | |
| | | | | | | VIS=0,7 | | |
| | | | | | | | VIS=0,8 | |
| | | | | | | | VIS=0,9 | |
| | | | | | | | VIS=1,0 | |
| | | | | | | | | VIS=1,0 |
| | | | | | | | | VIS=1,2 |

ВЫБОР ИСПЫТАТЕЛЬНЫХ МАТЕРИАЛОВ

| Задачи оценки | Используемый материал |
|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Общее качество со средним материалом | Обычный, "критичный, но не чрезмерно" |
| Пропускная способность, критичные применения (например, подача программ, пост-обработка и т. д.) | Набор, включающий очень критичный материал для тестируемых применений |
| Качество "адаптивных" систем | Материал, очень критичный для используемых "адаптивных" схем |
| Определение слабостей и возможных улучшений | Материал, критичный к определенным атрибутам |
| Определение причин, по которым системы выглядят различными | Широкий набор разнообразных материалов |
| Преобразование между различными стандартами | Критичный к различиям (например, скорости полей) |

СЕАНС ИСПЫТАНИЙ

- Продолжительность сеанса не должна превышать получаса. В начале первого сеанса следует провести примерно пять "тренировочных демонстраций для стабилизации мнения наблюдателей. Информация, полученная в ходе этих демонстраций, не должна учитываться в результатах испытания. Если необходимо проведение нескольких сеансов, то в начале каждого следующего сеанса необходимо проведение только трех "тренировочных демонстраций".

Структура представления сеанса испытания



ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ

Результаты должны быть представлены совместно со следующей информацией:

- подробное описание конфигурации испытания;
- подробное описание материалов для испытания;
- тип источника изображения и отображающих мониторов;
- количество и тип оценщиков;
- используемые эталонные системы;
- средняя оценка эксперимента;
- исходные и поправленные величины средней оценки и 95% доверительный интервал в том случае, если мнение одного или нескольких наблюдателей не принималось в расчет в соответствии с процедурой, описанной ниже.

МЕТОД *DSIS*

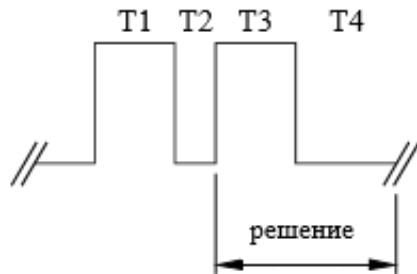
Метод DSIS (Double Stimulus Impairment Scale, Метод оценки искажений с двумя источниками воздействия). В ходе тестирования эксперту показывают сначала эталонное видео, а потом то же самое видео, подвергнутое ухудшающим воздействиям. После этого его просят оценить второй образец относительно первого по дискретной шкале. Во время сеанса, который длится до получаса, эксперту демонстрируют наборы изображений или их последовательности в случайном порядке и со случайными искажениями, охватывая все возможные комбинации. Неискаженное изображение также включается в последовательность, которая должна быть оценена.

| | | |
|--|---|-------------------------------------|
| | 5 | искажение не заметно |
| | 4 | искажение заметно, но не раздражает |
| | 3 | искажение слегка раздражает |
| | 2 | искажение раздражает |
| | 1 | искажение очень раздражает |

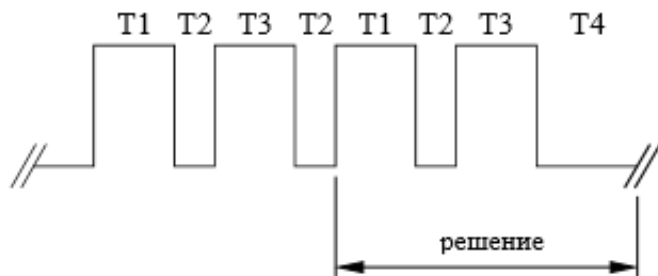
ОБЩАЯ КОМПОНОВКА ИСПЫТАТЕЛЬНОЙ СИСТЕМЫ ДЛЯ МЕТОДА DSIS



СТРУКТУРА ПРЕДСТАВЛЕНИЯ МАТЕРИАЛА ДЛЯ ИСПЫТАНИЙ



а) Вариант I



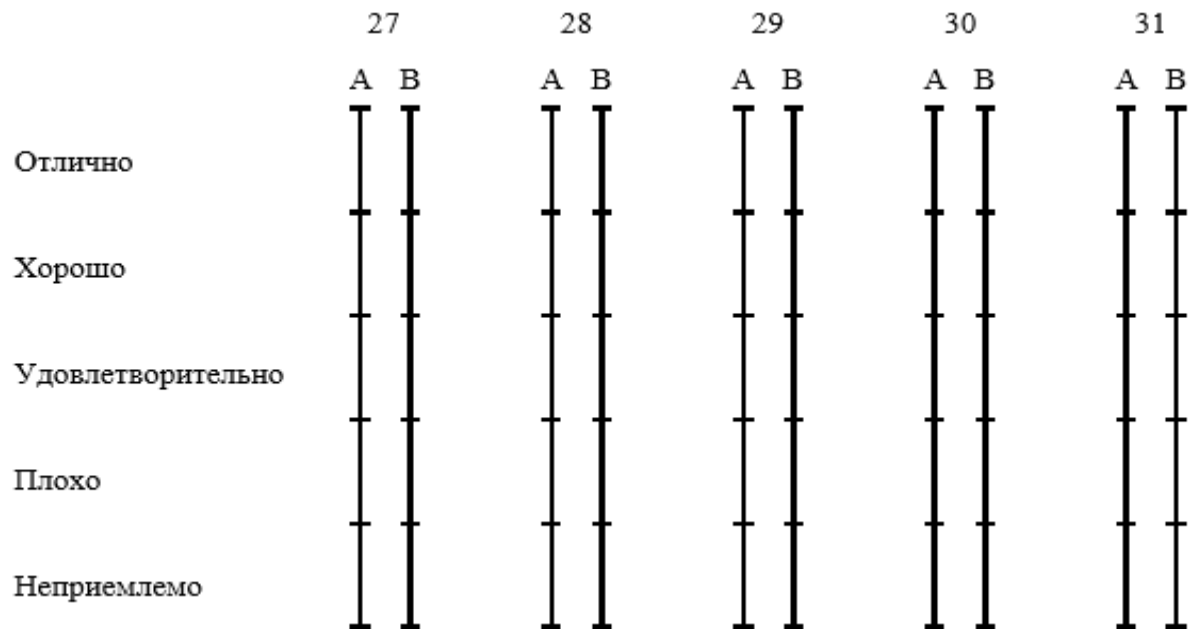
б) Вариант II

Этапы представления:

- $T_1 = 10$ с - Эталонное изображение
- $T_2 = 3$ с - Средний серый, создаваемый видеосигналом с уровнем примерно 200 мВ
- $T_3 = 10$ с - Испытательный сигнал
- $T_4 = 5-11$ с - Средний серый

МЕТОД DSCQS

Метод DSCQS (Double Stimulus Continuous Quality-Scale, Метод с двумя источниками воздействия и непрерывной шкалой качества и с использованием шкалы искажений). Наблюдателям предлагают просто оценить общее качество каждого представленного изображения, проставив отметку на вертикальной шкале. Такие вертикальные шкалы печатаются попарно для сдвоенного представления каждого тестируемого изображения. Эти шкалы представляют собой систему непрерывной оценки во избежание ошибок квантования, но они разделены на пять отрезков равной длины, соответствующих обычной пятибалльной шкале.



Метод DSCQS

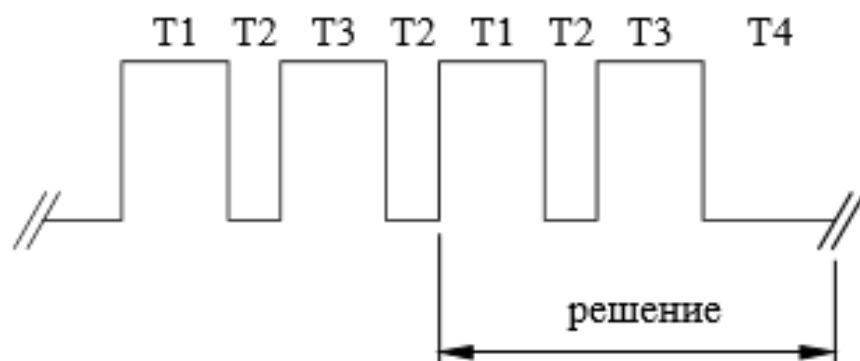
Наблюдателям предлагают просто оценить общее качество каждого представленного изображения, проставив отметку на вертикальной шкале

| | 27 | 28 | 29 | 30 | 31 |
|-------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | A B | A B | A B | A B | A B |
| Отлично | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Хорошо | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Удовлетворительно | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Плохо | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Неприемлемо | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

ОБЩАЯ КОМПОНОВКА ИСПЫТАТЕЛЬНОЙ СИСТЕМЫ ДЛЯ МЕТОДА *DSCQS*



СТРУКТУРА ПРЕДСТАВЛЕНИЯ МАТЕРИАЛА ДЛЯ ИСПЫТАНИЙ

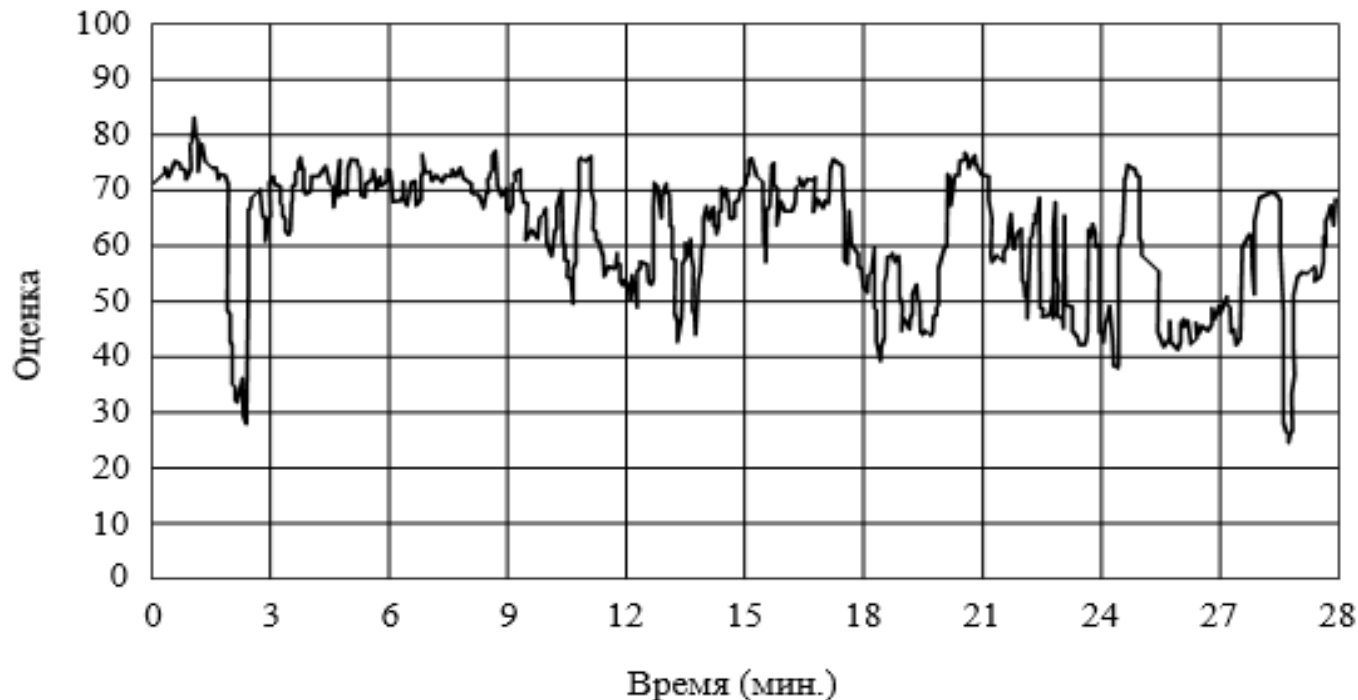


Фазы представления:

- | | |
|-------------|-----------------------------------------------------|
| T1 = 10 с | Тестовая последовательность А |
| T2 = 3 с | Средний серый, уровень видеосигнала примерно 200 мВ |
| T3 = 10 с | Испытательная последовательность В |
| T4 = 5–11 с | Средний серый |

МЕТОД SSCQE

SSCQE (Single Stimulus Continuous Quality Evaluation, Метод непрерывной оценки качества при одном источнике воздействия) - Непрерывно оценивает одну программу длительностью 10-20 минут, а не серию тестовых сцен. Оценка зависит в том числе и от содержания программы. Участники просматривают материал только один раз без эталонного источника.

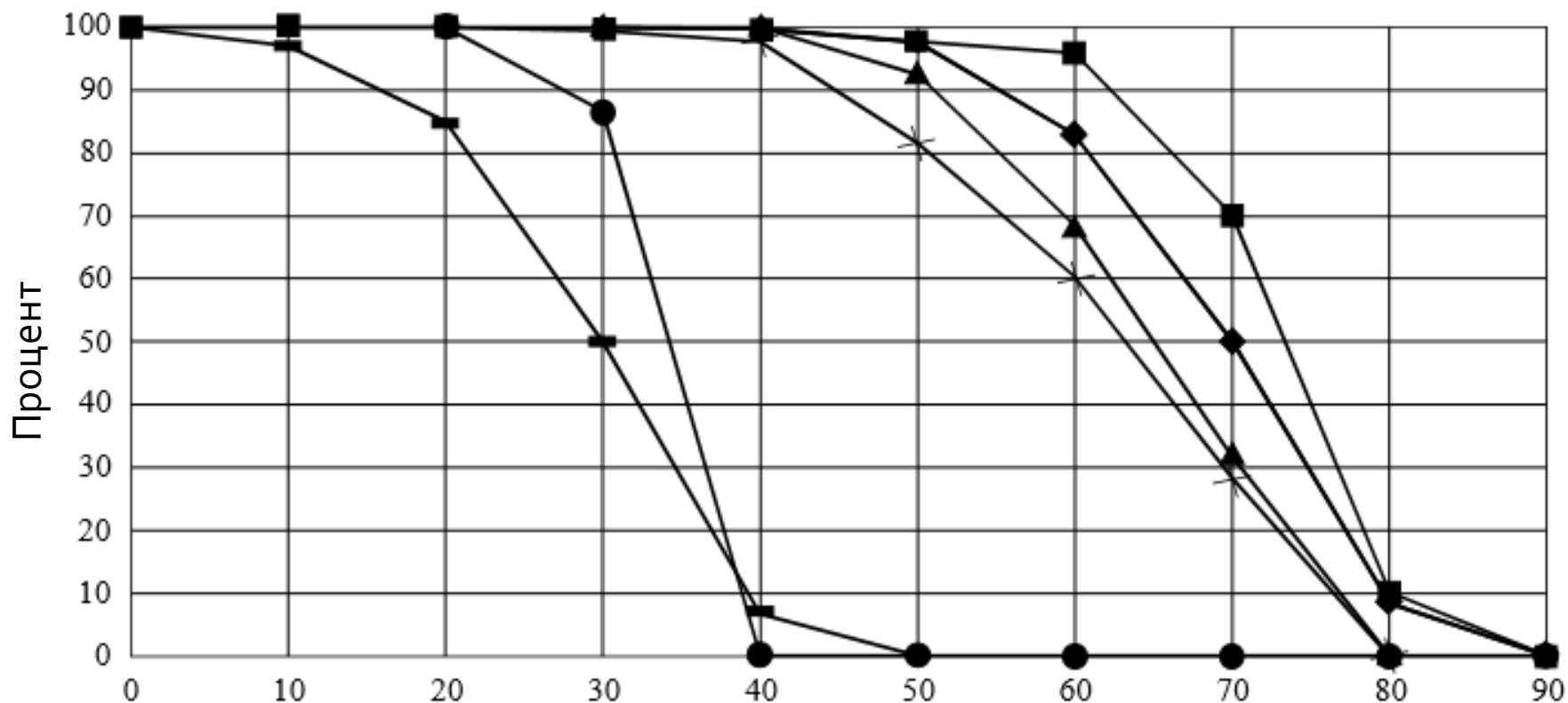


ОБЩАЯ ФОРМА ИСПЫТАТЕЛЬНОГО ПРОТОКОЛА

Участники должны присутствовать на сеансах испытаний следующего формата:

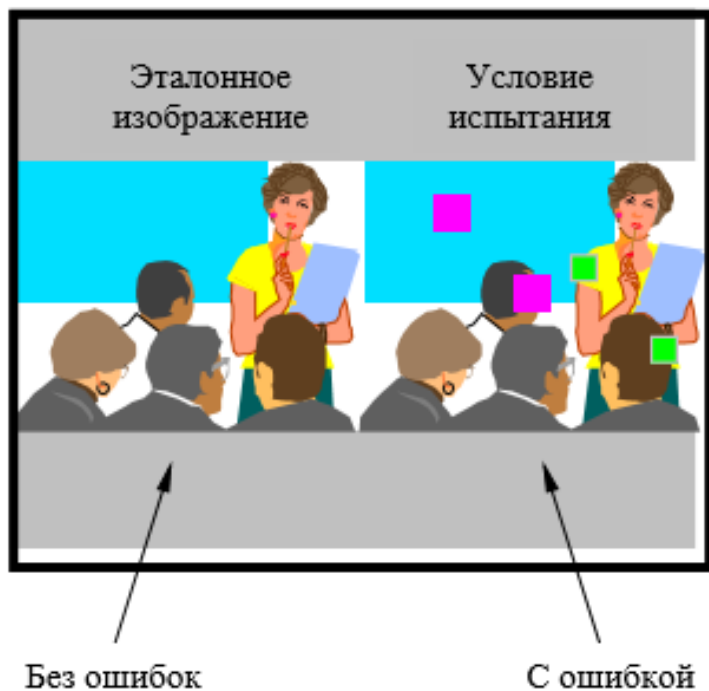
- – Фрагмент программы (ФП): ФП соответствует одному типу программы (например, спорт, новости, постановочная программа), обработанной в соответствии с одним из испытываемых параметров качества (ПК) (например, скорость в битах); каждый ФП должен длиться не менее 5 минут;
- – Сеанс испытаний (СИ): СИ – это последовательность одного или нескольких сочетаний ФП/ПК без разделения, имеющая псевдослучайный порядок. В каждом СИ хотя бы один раз имеют место все ФП и ПК, но необязательно все сочетания ФП/ПК; каждый СИ должен иметь продолжительность от 30 до 60 минут;
- – Демонстрация испытания (ДИ): ДИ – это испытание в полном объеме. ДИ может быть разделена на СИ, чтобы были удовлетворены требования в отношении максимальной продолжительности и с целью оценки качества по всем парам СП/ПК. Если количество пар СП/ПК ограничено, то ДИ может быть осуществлена путем повторения того же СИ для выполнения испытания в течение достаточного длительного периода времени.

СРЕДНИЕ ОЦЕНКИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ФРАГМЕНТА ПРОГРАММЫ Z, ПО КОТОРЫМ ПРОВОДИТСЯ ГОЛОСОВАНИЕ



Средние оценки последовательностей,
по которым проводится голосование

МЕТОД *SDSCE*



Метод SDSCE (Simultaneous Double Stimulus for Continuous Evaluation, Метод с двумя источниками одновременного воздействия для непрерывной оценки). Группа участников одновременно просматривает две последовательности: одну – эталонную, другую – относящуюся к одному из условий испытания. Участникам предлагается выявить различия между двумя последовательностями и оценить верность видеоинформации путем ручного устройства голосования. Участники проинформированы, на каком экране демонстрируется эталонное изображение, и им предлагается выразить свое мнение в процессе просмотра последовательностей на всем протяжении их демонстрации.

График зависимости оценок от времени видеопоследовательности "Go Pro"

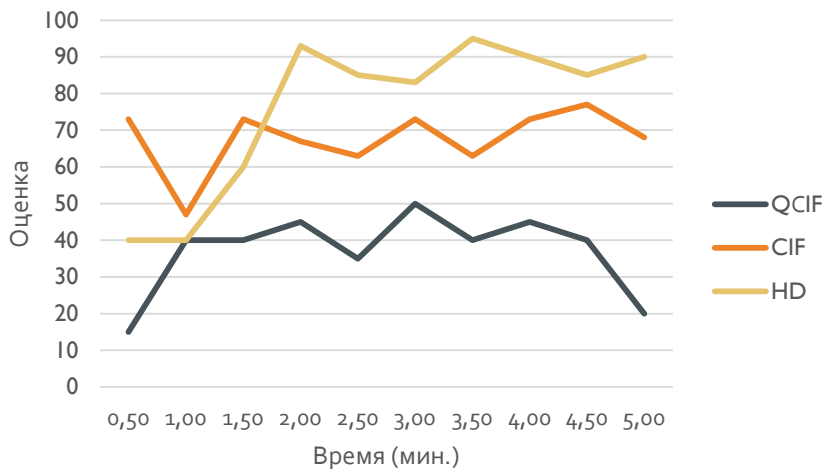


График зависимости оценок от времени видеопоследовательности "Hip Hop"

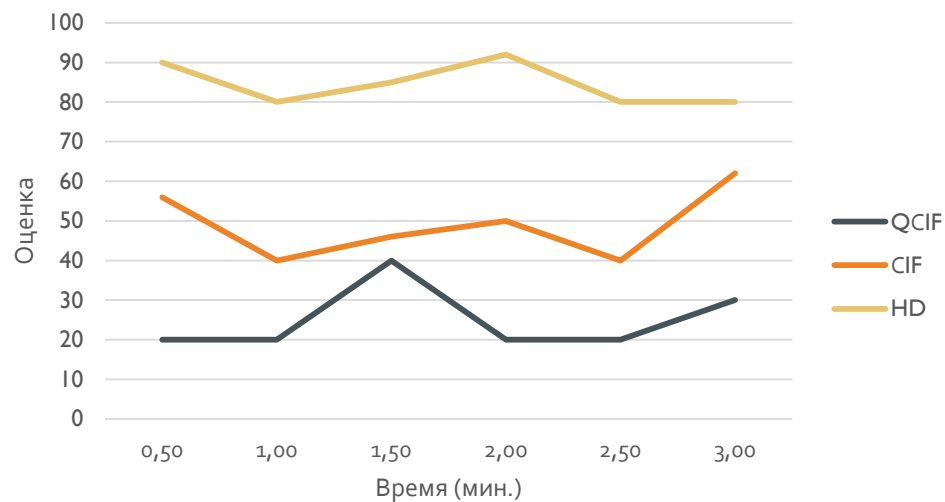


График зависимости оценки от времени видеопоследовательности "Опарыши"

