# ФРАКТАЛЬНЫЙ ПРЕПРОЦЕССИНГ КАДРОВ ДЛЯ УПРАВЛЕНИЯ СТРУКТУРОЙ GOP В СТАНДАРТИЗИРОВАННЫХ ВИДЕОКОДЕРАХ

*Пузий Анастасия Николаевна*

*Доцент кафедры «Системы телерадиовещания», Ташкентский университет информационных технологий им. Мухаммада аль-Хорезмий, Ташкент, Узбекистан*
*Email: puziy@tuit.uz*

*K E Y W O R D S*

H.266, группа кадров, смена сцены, фрактальная размерность, оптический поток, ориентация рёбер, препроцессинг видео

*A B S T R A C T*

Данная статья подчеркивает, что эффективность применения стандартизированных видеокодеров H.264, H.265 и H.266 зависит не только от перебора режимов внутри кадра, но и от того, насколько структура группы кадров (Group of Pictures, GOP) согласована со сменами сцен: ключевой кадр вне смены сцены расходует процессорную мощность кодера не рационально, а пропущенная смена сцены требует предсказывать кадр по несвязанному опорному. Практические детекторы смены сцены, основанные на статистике яркости, ложно срабатывают на вспышках, изменениях экспозиции и быстром движении камеры. В статье развивается предложенный автором метод фрактального препроцессинга кадров, который выполняется до начала работы кодера и не меняет синтаксис битового потока. Предложен алгоритм управления принятием решений кодера о начале группы кадров по пороговым значениям признаков с препроцессора GOP и локальном обновлении опорности, использующие изменение заполнения плиток рёбрами, расстояние Кульбака–Лейблера между гистограммами ориентаций рёбер и потерю надёжного оптического потока Лукаса–Канаде, а также отображение сигналов препроцессора на механизмы кодеков H.264, H.265, H.266 и AV1. Новизна состоит в использовании структурных признаков границ объектов совместно с надёжностью движения для различения смены сцены, поворота камеры, появления объекта и радиометрических помех. На тестовой последовательности из 360 кадров с 9 склейками предложенный алгоритм обнаружил все склейки без ложных срабатываний во всём рекомендуемом диапазоне порогов, тогда как гистограммный детектор при наилучшем пороге дал 2 ложных срабатывания на вспышках.

## 1. ВВЕДЕНИЕ

Стандарты H.264, H.265 и H.266 по мере своего появления снижали битовую скорость примерно вдвое при сохранении визуального качества [1-5], однако ценой многократного роста вычислительной сложности: кодер H.266 сложнее кодера H.265 примерно в 5 раз в конфигурации Low-Delay и в 31 раз в конфигурации All-Intra [6, 7]. Основная часть затрат приходится на перебор разбиений блоков по схеме QTMT и на оценку движения [8, 9].

В предшествующей работе автора [10] предложен метод вспомогательного фрактального препроцессинга, который до запуска стандартного кодера оценивает фрактальную размерность границ объектов, плотность рёбер и надёжность оптического потока и передаёт кодеру подсказки о глубине разбиения, квантовании и начальных векторах движения без изменения синтаксиса битового потока. На модели внутрикадрового кодера H.266 подсказки по произведению размерности и плотности рёбер сократили число проверок RDO на 35,2 % при росте BD-rate на 1,01 %; фрактальная размерность оказалась в 12–15 раз устойчивее дисперсии и энтропии к изменению контраста, а ограничение слежения за точками областями интереса сократило время расчёта оптического потока на 24 % [10]. Решения о структуре группы кадров, также предусмотренные схемой препроцессора, в [10] не рассматривались.

Между тем структура группы кадров (GOP) существенно влияет на эффективность кодирования. Ключевой кадр, поставленный

(cc) BY 

вне смены сцены, расходует процессорную мощность кодера не рационально, а пропущенная смена сцены требует предсказывать кадр по несвязанному опорному кадру (I-кадру). Задача обнаружения границ сцен изучается с начала 1990-х годов.: гистограммные методы сравнивают распределения яркости соседних кадров [11], методы на основе рёбер фиксируют появление новых контуров вдали от прежних [12]. Систематизация подходов и результаты сравнительных испытаний показывают, что основными источниками ошибок остаются вспышки, изменения освещения и быстрое движение камеры [13, 14]. В видеокодерах решение о вставке ключевого кадра часто принимается по простым пороговым признакам, чувствительным к тем же помехам.

Признаки, которые препроцессор уже вычисляет для управления разбиением - это геометрия рёбер и надёжность движения. Они по своей природе устойчивы к радиометрическим помехам: фрактальное описание сохраняется при линейных преобразованиях яркости [15], а контуры описывают структуру сцены, а не её освещённость [12, 16]. Это делает их естественной основой и для решений о GOP.

Цель исследования – разработать алгоритм управления структурой GOP и локальным обновлением опорности стандартизированных кодеров по сигналам фрактального препроцессора и оценить их точность. Для достижения цели решались следующие задачи:

1. определить набор сигналов препроцессора и их отображение на механизмы кодеков H.264, H.265, H.266 и AV1;
2. сформулировать детерминированный алгоритм начала GOP, мягкого обновления и локального обновления опорности;
3. экспериментально проверить поведение признаков при различных событиях в кадре и сравнить точность обнаружения смен сцен с классическими детекторами.

Эксперименты проводились на синтезированной видеопоследовательности из 360 кадров с известными моментами смены сцен, построенной из 12 тестовых изображений пакета scikit-image [17].

## 2. ТЕОРЕТИЧЕСКАЯ ОСНОВА

### 2.1. Фрактальный препроцессор

Структурная схема препроцессора приведена на рис. 1; подробное описание его блоков и оценка подсказок по разбиению даны в [10]. Препроцессор работает со скользящим окном из трёх кадров по яркостной компоненте Y. Операторы Собела и Кэнни строят карту рёбер [18, 19], для плиток 32×32 оценивается коробочная фрактальная размерность D как наклон зависимости log N(ε) от log(1/ε) по масштабам ε = 2, 4, 8, 16 и т.д.[16, 20]:

$$D \approx \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})}{\sum_n (x_n - \bar{x})^2}, \; x_n = \log \frac{1}{\varepsilon_n},$$

$$y_n = \log N(\varepsilon_n) \qquad (1)$$

где N(ε) – число ячеек сетки со стороной ε, содержащих пиксели рёбер. В значимых областях выбираются точки Ши–Томаси [21] и отслеживаются пирамидальным методом Лукаса–Канаде [22, 23] с помощью функций библиотеки OpenCV [24, 25]. Вектор считается валидным, если точка, отслеженная вперёд и назад, возвращается в исходное положение с ошибкой менее 1 пикселя.

(cc) BY

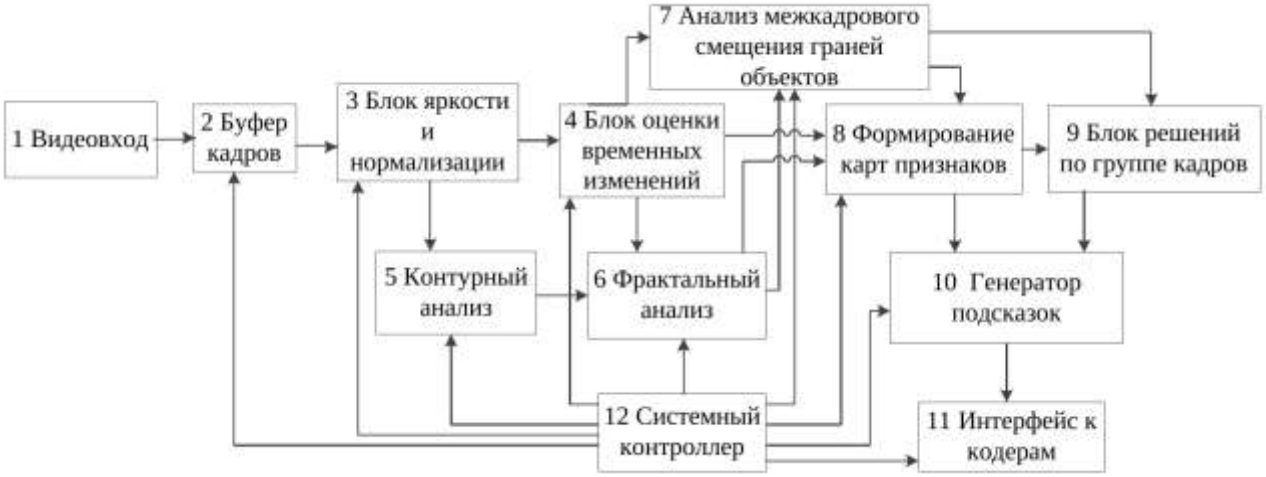



***Рис. 1. Блок-схема метода адаптации к содержимому кадров на основе вычисления фрактальной размерности элементов изображения***

Решение о структуре GOP принимает блок 9 по признакам, формируемым блоками 4–8, а генератор подсказок (блок 10) и интерфейс к кодерам (блок 11) передают эти решения конкретному кодеру.

### 2.2. Структура группы кадров в стандартизированных кодеках

Группа кадров начинается с кадра, декодируемого без ссылок на предыдущие. В H.264 это IDR-кадр [4]; в H.265 и H.266, помимо IDR, введён кадр CRA, допускающий «открытую» GOP, при которой следующие за ним кадры могут ссылаться на кадры предыдущей группы [2, 5]. В H.266 стандартизирован также постепенный переход к декодированию (GDR), при котором опорность восстанавливается по частям кадра [2]; в AV1 аналогичные функции выполняют ключевые кадры и кадры только с внутренним предсказанием [26]. Поправки параметра квантования на уровне блока предусмотрены во всех этих стандартах.

### 2.3. Признаки смены сцены

Для решений о GOP препроцессор использует два структурных признака рёбер и признаки надёжности движения. Первый – изменение заполнения плиток рёбрами между соседними кадрами:

$$\Delta E_t = \frac{1}{K_T}\sum_{k=1}^{K_T}\left|\rho_k(t) - \rho(t-1)\right| \qquad (2)$$

где $\rho_k(t)$ - плотность рёбер в плитке k кадра t, $K_T$ – число плиток. Второй признак - это расстояние Кульбака-Лейблера между гистограммами ориентаций рёбер $P_{k,t}$ плитки в текущем и предыдущем кадрах [27]:

$$KL_{k,t} = \sum_{b=1}^{8} P_{k,t}(b)\ln\frac{P_{k,t}(b)}{P_{k,t-1}(b)} \qquad (3)$$

где гистограммы строятся по 8 интервалам направления градиента $\theta=\text{arctg}(G_y/G_x)$, взвешиваются модулем градиента и регуляризуются малой добавкой, исключающей нулевые интервалы. Для кадра используется медиана $KL_t$ по текстурированным плиткам: её превышение порогом означает, что ориентации перестроились в большинстве плиток. Признаки движения – число валидных векторов $N_{LK}$, число точек-кандидатов $N_{cand}$ и доля $g_t$ текстурированных плиток, в которых ни одна точка не отслежена надёжно (доля плиток без потока).

## 3. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

### 3.1. Сигналы препроцессора и алгоритм управления кодером

Препроцессор формирует следующие карты и события, передаваемые кодеру по «побочным каналам» (side channels) параллельно основной обработке видео; они также записываются в журнал диагностики и не задерживают работу кодера:

(cc) BY 

- - dqp map - поправки параметра квантования по плиткам: в областях с фракталоподобной структурой, прослеженных во времени, QP уменьшается, вне их - увеличивается;
- -refresh map – локальный сигнал обновления опорности в плитках, где накопились ошибки предсказания, о чём говорят отсутствие согласованных векторов и распад контуров; это аналог локального intra-обновления стандартизированных кодеков;
- - motion gate tiles – маска плиток без стабильного оптического потока, в которых сложные межкадровые режимы нецелесообразны и приоритет отдаётся внутрикадровому предсказанию;
- - события GOP $Q_{gop}$ = (i, IDR/CRA/continue), где i – номер кадра, а второй элемент - это тип события; например, (120, continue) означает, что на кадре 120 действий по GOP не требуется.

Решения о GOP принимаются по детерминированному алгоритму (рис. 2), где правило смены сцены: ключевой кадр IDR назначается, если одновременно выполняются условия:

$$\Delta E_t > \tau_1 \; u \; \underset{k}{med}\, KL_{k,t} > \tau_2 \qquad (4)$$

Совместное использование двух признаков необходимо потому, что по отдельности каждый из них реагирует и на события, не требующие нового ключевого кадра: поворот камеры меняет ориентации при почти неизменном заполнении рёбрами, а шум и экспозиция слабо меняют заполнение, не затрагивая ориентаций. Еще одно условие потери движения, при котором в группе кадров назначается IDR следующее:

$$N_{LK} \le N_{\min}, \;\; N_{cand} \ge 50, \;\; F_{edge} \ge E_{\min} \qquad (5)$$

где $F_{edge}$ – доля плиток, содержащих рёбра, $N_{min} = 2$, $E_{min} = 0{,}1$. Если текстура в кадре есть, а надёжных векторов между соседними кадрами практически нет, межкадровое предсказание теряет смысл, и участок следует отделить ключевым кадром. Мягкое правило алгоритма: при доле плиток без потока $g_t \ge 0{,}6$ планируется кадр CRA на ближайших кадрах. После каждого срабатывания вводится задержка (cool-down) в 8–16 кадров, исключающая повторные срабатывания из-за шума.

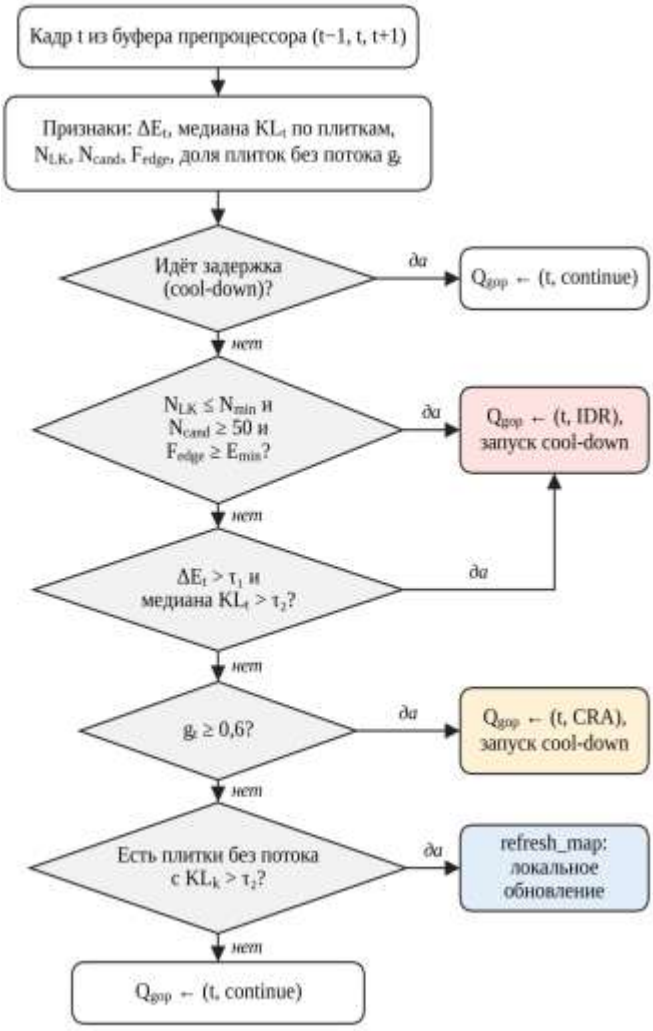



***Рис. 2. Алгоритм принятия решения о начале GOP по признакам препроцессора***

В алгоритме правило локального обновления в кадрах: если в отдельных плитках нет стабильных векторов и ориентации рёбер перестроились ($KL_{k,t}>\tau_2$), но глобальные правила не сработали, эти плитки помечаются в refresh map для локального обновления опорности без генерации полного опорного I-кадра. Правила для областей интереса (Regions of Interest, ROI) дополняют подсказки по разбиению из [10]: в плитках с низкой плотностью рёбер и без существенного движения понижается приоритет тонкого разбиения и расширенного поиска движения, а в плитках ROI разрешаются более сложные режимы, в том числе геометрические

(cc) BY 

разбиения и аффинные модели движения, причём доминирующее направление рёбер служит подсказкой при выборе направленных предсказаний.

Все подсказки выражаются через механизмы, уже предусмотренные стандартами или относящиеся к решениям кодера (табл. 1), поэтому синтаксис битового потока не меняется. Адаптер подсказок формирует для кодера файл параметров квантования и типов кадров (для H.264/H.265), список ключевых кадров (для AV1) или разбивает последовательность на независимо кодируемые отрезки (для H.266).

***Таблица 1.***

***Отображение сигналов препроцессора на механизмы стандартизированных кодеков***

| **Сигнал препроцессора** | **H.264** | **H.265** | **H.266** | **AV1** |
|---|---|---|---|---|
| Событие IDR | IDR-кадр | IDR-кадр | IDR-кадр | Ключевой кадр |
| Событие CRA | I-кадр с SEI точки восстановления | CRA-кадр | CRA-кадр | Кадр только с внутренним предсказанием |
| refresh map | Принудительные intra-макроблоки | Принудительные intra-блоки | GDR или принудительные intra-блоки | Принудительные intra-блоки |
| dqp map | mb qp delta | cu qp delta | cu qp delta | delta q суперблока |
| motion gate tiles | Ограничение режимов и области поиска в кодере (без изменения синтаксиса) | | | |

### 3.2. Методика эксперимента

Для проверки предложенного алгоритма из 12 изображений пакета scikit-image [17] была синтезирована последовательность из 360 кадров размером 256×384: 12 сцен по 30 кадров с панорамированием камеры, 9 резких склеек и 2 наплыва длительностью 8 кадров. В каждую сцену с 14-го кадра внесена одна помеха, не требующая нового ключевого кадра: вспышка (+60 уровней яркости на 2 кадра), плавное снижение экспозиции до 55 %, аддитивный шум $\sigma = 6$, плавный поворот камеры на 1,5° за кадр, появление движущегося объекта или быстрое панорамирование (10 пикселей за кадр). Реакция признаков на резкий поворот камеры оценивалась отдельно: каждый кадр сравнивался со своей копией, повёрнутой на 10° и 20°.

Алгоритм применялся с фиксированными порогами из диапазонов $\tau_1 = 0{,}06–0{,}12$ и $\tau_2 = 0{,}35–0{,}45$ без подстройки на этой последовательности; дополнительно проверялось значение $\tau_2 = 0{,}25$. Для сравнения использовались классические детекторы - разность гистограмм яркости [11] и средняя абсолютная разность кадров с порогом, наилучшим для них на этой же последовательности, что заведомо даёт им преимущество. Обнаружение засчитывалось при совпадении со склейкой с точностью ±1 кадр; срабатывание внутри наплыва считалось обнаружением перехода; для всех детекторов действовала одинаковая задержка в 8 кадров.

## 4. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

### 4.1. Поведение признаков при различных событиях в кадре

В табл. 2 для каждого типа событий приведены ожидаемое поведение признаков и измеренные медианные значения $\Delta E_t$, медианы $KL_t$ и доли плиток без потока $g_t$.

(cc) BY 

***Таблица 2.***

***Совместное поведение признаков заполнения рёбрами (порог $\tau_1$) и их ориентации (порог $\tau_2$) при различных событиях в кадре***

| **Событие в кадре** | **ΔE ($\tau_1$), ожидание** | **KL ($\tau_2$), ожидание** | **ΔE, изм.** | **KL, изм.** | **g, изм.** | **Значение для кодера** |
|---|---|---|---|---|---|---|
| Обычное движение камеры | – | – | 0,024 | 0,027 | 0,00 | Продолжение GOP |
| Изменение экспозиции | Слабо растёт | ≈ 0 | 0,018 | 0,029 | 0,00 | Нет структурной смены |
| Шум | Слабо растёт | ≈ 0 | 0,014 | 0,036 | 0,01 | Нет структурной смены |
| Вспышка | Слабо растёт | ≈ 0 | 0,028 | 0,022 | 0,03 | Нет структурной смены |
| Плавный поворот, 1,5° за кадр | Постоянен | Растёт | 0,030 | 0,021 | 0,00 | Поток отслеживает поворот |
| Резкий поворот, 10° / 20° | Постоянен | Растёт | 0,056 / 0,082 | 0,295 / 0,502 | – | Меняется ориентация кадра |
| Появление объекта | Растёт | Растёт | 0,023 | 0,053 | 0,09 | Локальная новая структура |
| Быстрое панорамирование | – | – | 0,051 | 0,106 | 0,00 | Продолжение GOP |
| Наплыв | – | – | 0,035 | 0,063 | 0,02 | Плавный переход |
| Смена сцены (склейка) | Сильно растёт | Сильно растёт | 0,190 | 0,903 | 1,00 | Полная замена структуры |

Измерения подтверждают ожидаемое поведение признаков. При смене сцены медиана $\Delta E_t$ (0,190) в 8 раз, а медиана $KL_t$ (0,903) в 33 раза превышают значения при обычном движении, и почти все плитки теряют поток ($g_t = 1$). Изменения экспозиции, шум и вспышки почти не меняют ни заполнения рёбрами, ни ориентаций, что подтверждает структурный характер признаков. Резкий поворот камеры увеличивает KL в 11–19 раз при умеренном росте ΔE, как и предполагалось; при плавном повороте изменение между соседними кадрами мало, а поток его уверенно отслеживает. Появление объекта даёт лишь локальный рост признаков, поэтому глобальные правила на него не реагируют, а обрабатывает его правило локального обновления.

### 4.2. Точность обнаружения смен сцен

(cc) BY

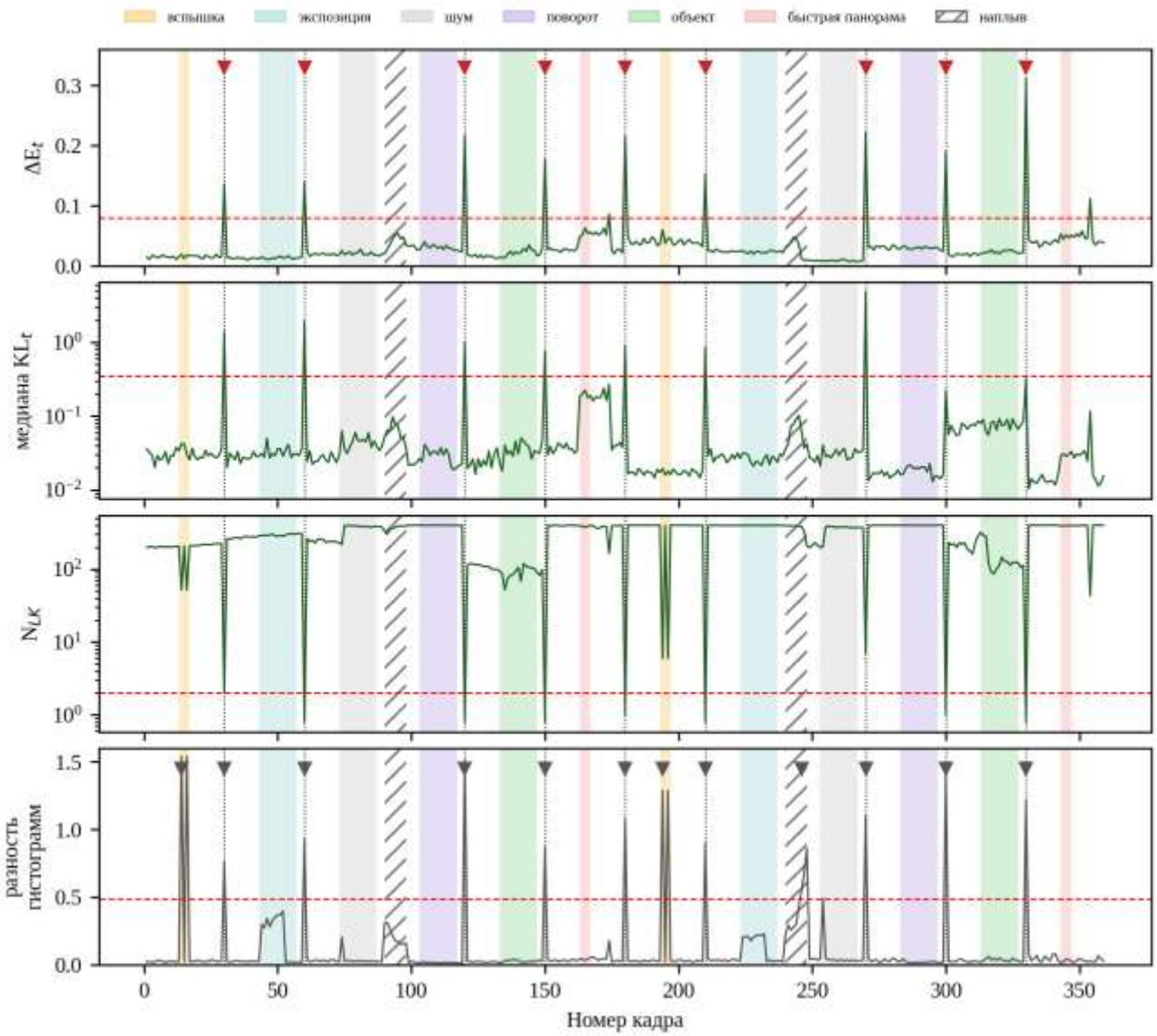



***Рис. 3. Признаки препроцессора и разность гистограмм яркости на тестовой последовательности: пунктирные вертикали – склейки, заливка – помехи, штриховка – наплывы, горизонтальные линии – пороги, треугольники – срабатывания детекторов***

На рис. 3 представлены результаты анализа чувствительности признаков, предложенных автором и разности гистограмм (стандартизированного кодера) в качестве параметров, на которые может ссылаться кодер при принятии решений о начале новой группы кадров в процессе межкадрового кодирования. Здесь пунктирными вертикальными линиями представлены склейки, заливкой - помехи, штриховкой - наплывы, горизонтальные линии представляют собой пороги, а треугольники показывают моменты срабатывания детекторов.

***Таблица 3.***

***Обнаружение смен сцен на тестовой последовательности (9 склеек, 2 наплыва)***

| Детектор | Найдено склеек | Ложных срабатываний | Источник ложных срабатываний | Найдено наплывов |
|---|---|---|---|---|
| Предлагаемые правила, $\tau_1$ = 0,06–0,12, $\tau_2$ = 0,35–0,45 | 9 | 0 | – | 0 |

(cc) BY *dtai.uz*

| Детектор | Найдено склеек | Ложных срабатываний | Источник ложных срабатываний | Найдено наплывов |
|---|---|---|---|---|
| То же, $\tau_1 = 0{,}06–0{,}08$, $\tau_2 = 0{,}25$ | 8 | 1 | Движение камеры | 0 |
| Только правило смены сцены ($\tau_1 = 0{,}08$, $\tau_2 = 0{,}35$) | 7 | 0 | – | 0 |
| Только правило потери движения | 8 | 0 | – | 0 |
| Мягкое правило ($g \geq 0{,}6$) | 9 | 2 | Вспышка, движение | 0 |
| Разность гистограмм яркости (лучший порог) | 9 | 2 | Вспышки | 1 |
| Средняя разность кадров (лучший порог) | 6 | 0 | – | 0 |
| Средняя разность кадров (порог полного обнаружения) | 8 | 5 | Вспышки, движение | 0 |

Сочетание правил смены сцены и потери движения обнаружило все 9 склеек без ложных срабатываний во всём рекомендуемом диапазоне порогов (табл. 3, рис. 3). Правила дополняют друг друга: две склейки между сценами с похожей текстурой (brick и gravel, moon и grass) дали медиану $KL_t$ ниже $\tau_2$ и были обнаружены по потере движения, а одна склейка, на которой сохранилось 7 случайно совпавших векторов, - по правилу смены сцены. Снижение $\tau_2$ до 0,25 привело к ложному срабатыванию на движении камеры, после которого задержка подавила обнаружение ближайшей склейки; это подтверждает целесообразность диапазона $\tau_2 \approx 0{,}35–0{,}45$.

Гистограммный детектор даже при наилучшем для него пороге ложно срабатывает на обеих вспышках, а средняя разность кадров не позволяет одновременно обнаружить все склейки и избежать ложных срабатываний на вспышках и быстром движении. Мягкое правило CRA, как и ожидалось, чувствительнее: оно срабатывает на кратковременной потере потока при вспышке, что допустимо для открытой GOP. Наплывы ни один из детекторов, кроме гистограммного в одном случае, не обнаружил.

### 4.3. Локальное обновление и вычислительные затраты

В проведенных экспериментах с предложенным алгоритмом правило локального обновления при появлении объекта пометило в refresh map в среднем 0,43 плитки на кадр, из них 62 % были на самом объекте, преимущественно на его переднем крае; в кадрах с обычным движением - 0,13 плитки на кадр из 96. Таким образом алгоритм реагирует на локальную новую структуру, но с небольшой полнотой.

Расчёт признаков GOP вместе с отслеживанием точек в однопоточной реализации на Python занял в среднем 18 мс на кадр 256×384. Эти признаки строятся из тех же карт рёбер и векторов, что и подсказки по разбиению, поэтому их добавление почти не увеличивает затраты препроцессора, условия окупаемости которого рассмотрены в [10].

## 5. АНАЛИЗ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Устойчивость предложенного алгоритма к радиометрическим помехам объясняется структурной природой признаков. Гистограммный подход сравнивает распределения яркости и потому реагирует на

(cc) BY 

вспышки и изменения экспозиции [11, 13]; признаки рёбер, как и в методе [12], описывают геометрию сцены и на такие помехи почти не реагируют. Новым по сравнению с [12] является сочетание статистики ориентаций рёбер с надёжностью оптического потока: оно позволяет различать смену сцены, поворот камеры и появление объекта (табл. 2) и направлять кодеру разные действия – IDR, CRA или локальное обновление. Этот результат согласуется с данными [10], где те же структурные признаки оказались в 12–15 раз устойчивее дисперсии и энтропии к изменению контраста.

Слабым местом, общим для всех проверенных детекторов, остаются плавные переходы, для обнаружения которых требуется анализ признаков на более длинном окне.

## 6. ЗАКЛЮЧЕНИЕ

В работе метод фрактального препроцессинга кадров, ранее применённый для сокращения перебора разбиений блоков, расширен на управление структурой группы кадров и локальным обновлением опорности стандартизированных кодеров. Определён набор сигналов препроцессора: поправки квантования, карта локального обновления, маска плиток без стабильного движения и события GOP, и их влияние на механизмы кодеков H.264, H.265, H.266 и AV1, не требующее изменения синтаксиса битового потока.

Предложенный алгоритм учитывает изменение заполнения рёбрами, перестройку их ориентаций и потерю надёжного движения. Эксперимент подтвердил ожидаемое поведение признаков: при смене сцены они возрастают в 8–33 раза, тогда как вспышки, изменения экспозиции и шум их почти не меняют, а поворот камеры и появление объекта различимы по сочетанию признаков. Алгоритм обнаружил все смены сцен на тестовой последовательности без ложных срабатываний в широком диапазоне порогов, превзойдя гистограммный детектор и детектор разности кадров, при затратах около 18 мс на кадр.

Рекомендуется принимать решения о начале GOP по совокупности структурных признаков и надёжности движения, используя пороги $\tau_1 \approx 0{,}06–0{,}12$ и $\tau_2 \approx 0{,}35–0{,}45$.

### Список литературы

1. ITU-T Recommendation H.266 (08/2020). Versatile video coding. Geneva: International Telecommunication Union, 2020. URL: https://www.itu.int/rec/T-REC-H.266 (дата обращения: 26.09.2026).
2. Bross B., Wang Y.-K., Ye Y. [et al.] Overview of the Versatile Video Coding (VVC) standard and its applications // IEEE Transactions on Circuits and Systems for Video Technology. 2021. Vol. 31, no. 10. P. 3736–3764. DOI: 10.1109/TCSVT.2021.3101953.
3. Bross B., Chen J., Ohm J.-R. [et al.] Developments in international video coding standardization after AVC, with an overview of Versatile Video Coding (VVC) // Proceedings of the IEEE. 2021. Vol. 109, no. 9. P. 1463–1493. DOI: 10.1109/JPROC.2020.3043399.
4. Wiegand T., Sullivan G.J., Bjøntegaard G. [et al.] Overview of the H.264/AVC video coding standard // IEEE Transactions on Circuits and Systems for Video Technology. 2003. Vol. 13, no. 7. P. 560–576.
5. Sullivan G.J., Ohm J.-R., Han W.-J. [et al.] Overview of the High Efficiency Video Coding (HEVC) standard // IEEE Transactions on Circuits and Systems for Video Technology. 2012. Vol. 22, no. 12. P. 1649–1668.
6. Pakdaman F., Adelimanesh M.A., Gabbouj M. [et al.] Complexity analysis of next-generation VVC encoding and decoding // 2020 IEEE International Conference on Image Processing (ICIP). IEEE, 2020. P. 3134–3138.
7. Bossen F., Sühring K., Wieckowski A. [et al.] VVC complexity and software implementation analysis // IEEE Transactions on Circuits and Systems for Video Technology. 2021. Vol. 31, no. 10. P. 3765–3778. DOI: 10.1109/TCSVT.2021.3072204.
8. Li T., Xu M., Tang R. [et al.] DeepQTMT: a deep learning approach for fast QTMT-based CU partition of intra-mode VVC // IEEE Transactions on Image Processing. 2021. Vol. 30. P. 5377–5390.

(cc) BY 

9. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. М.: Техносфера, 2005. 368 с.
10. Puziy A. Fractal frame preprocessing to accelerate video encoding by standardized encoders // International Journal of Advanced Research in Science, Engineering and Technology. 2026. Vol. 13, no. 9. P. 25472–25481. URL: https://www.ijarset.com/upload/2026/september/12-puziy-18.pdf (дата обращения: 30.09.2026).
11. Zhang H.J., Kankanhalli A., Smoliar S.W. Automatic partitioning of full-motion video // Multimedia Systems. 1993. Vol. 1, no. 1. P. 10–28. DOI: 10.1007/BF01210504.
12. Zabih R., Miller J., Mai K. A feature-based algorithm for detecting and classifying scene breaks // Proceedings of the 3rd ACM International Conference on Multimedia. San Francisco, 1995. P. 189–200. DOI: 10.1145/217279.215266.
13. Hanjalic A. Shot-boundary detection: unraveled and resolved? // IEEE Transactions on Circuits and Systems for Video Technology. 2002. Vol. 12, no. 2. P. 90–105.
14. Smeaton A.F., Over P., Doherty A.R. Video shot boundary detection: seven years of TRECVid activity // Computer Vision and Image Understanding. 2010. Vol. 114, no. 4. P. 411–418. DOI: 10.1016/j.cviu.2009.03.011.
15. Pentland A.P. Fractal-based description of natural scenes // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984. Vol. PAMI-6, no. 6. P. 661–674. DOI: 10.1109/TPAMI.1984.4767591.
16. Mandelbrot B.B. The fractal geometry of nature. San Francisco: W.H. Freeman, 1982. 468 p.
17. van der Walt S., Schönberger J.L., Nunez-Iglesias J. [et al.] scikit-image: image processing in Python // PeerJ. 2014. Vol. 2. Art. e453. DOI: 10.7717/peerj.453.
18. Canny J. A computational approach to edge detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986. Vol. PAMI-8, no. 6. P. 679–698.
19. Гонсалес Р., Вудс Р. Цифровая обработка изображений. 3-е изд. М.: Техносфера, 2012. 1104 с.
20. Sarkar N., Chaudhuri B.B. An efficient differential box-counting approach to compute fractal dimension of image // IEEE Transactions on Systems, Man, and Cybernetics. 1994. Vol. 24, no. 1. P. 115–120.
21. Shi J., Tomasi C. Good features to track // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 1994. P. 593–600.
22. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI). Vancouver, 1981. P. 674–679.
23. Bouguet J.-Y. Pyramidal implementation of the Lucas Kanade feature tracker: description of the algorithm. Intel Corporation, Microprocessor Research Labs, 2000.
24. OpenCV: Optical Flow [Электронный ресурс]. URL: https://docs.opencv.org/4.x/d4/dee/tutorial_optical_flow.html (дата обращения: 26.09.2026).
25. Bradski G. The OpenCV Library // Dr. Dobb's Journal of Software Tools. 2000. Vol. 25, no. 11. P. 120–125.
26. Chen Y., Mukherjee D., Han J. [et al.] An overview of core coding tools in the AV1 video codec // 2018 Picture Coding Symposium (PCS). San Francisco: IEEE, 2018. P. 41–45.
27. Kullback S., Leibler R.A. On information and sufficiency // The Annals of Mathematical Statistics. 1951. Vol. 22, no. 1. P. 79–86.

(cc) BY 