

ВЕСЕЛОВ Антон Игоревич

ОБРАБОТКА ВИДЕОИНФОРМАЦИИ В СИСТЕМАХ СЖАТИЯ, ОСНОВАННЫХ НА ПРИНЦИПАХ КОДИРОВАНИЯ ЗАВИСИМЫХ ИСТОЧНИКОВ

Специальность 05.12.13 – Системы, сети и устройства
телекоммуникаций

Научный руководитель: д-р. техн. наук, доцент
Тюрликов Андрей Михайлович

2016

Актуальность работы

- Современные приложения с ограничениями на сложность кодера:
 - медицинские датчики;
 - мобильные источники видеоинформации;
 - комплексы аппаратуры космического базирования;
 - и т.д.

Актуальность работы

- Современные приложения с ограничениями на сложность кодера:
 - медицинские датчики;
 - мобильные источники видеоинформации;
 - комплексы аппаратуры космического базирования;
 - и т.д.
- Существующие методы сжатия видеоданных

	H.264 (режим Inter⁽¹⁾)
Эффективность сжатия	↑
Сложность кодера	↑

⁽¹⁾ Кодирование межкадровой разности, высокая сложность.

Актуальность работы

- Современные приложения с ограничениями на сложность кодера:
 - медицинские датчики;
 - мобильные источники видеоинформации;
 - комплексы аппаратуры космического базирования;
 - и т.д.
- Существующие методы сжатия видеоданных

	H.264 (режим Inter⁽¹⁾)	H.264 (режим Intra⁽²⁾)
Эффективность сжатия	↑	↓
Сложность кодера	↑	↓

⁽¹⁾ Кодирование межкадровой разности, высокая сложность.

⁽²⁾ Кодирование кадров независимо, низкая сложность.

Актуальность работы

- Современные приложения с ограничениями на сложность кодера:
 - медицинские датчики;
 - мобильные источники видеоинформации;
 - комплексы аппаратуры космического базирования;
 - и т.д.
- Существующие методы сжатия видеоданных

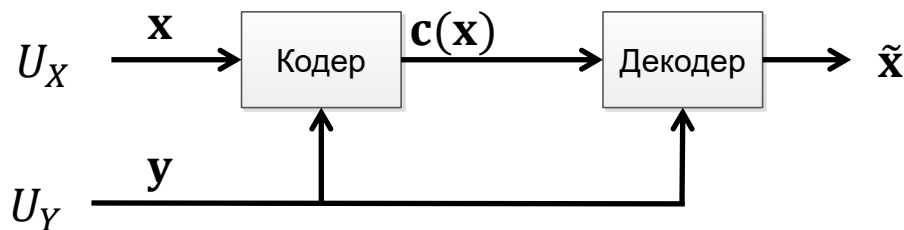
	H.264 (режим Inter⁽¹⁾)	H.264 (режим Intra⁽²⁾)	Распределенное кодирование
Эффективность сжатия	↑	↓	↑
Сложность кодера	↑	↓	↓

(1) Кодирование межкадровой разности, высокая сложность.

(2) Кодирование кадров независимо, низкая сложность.

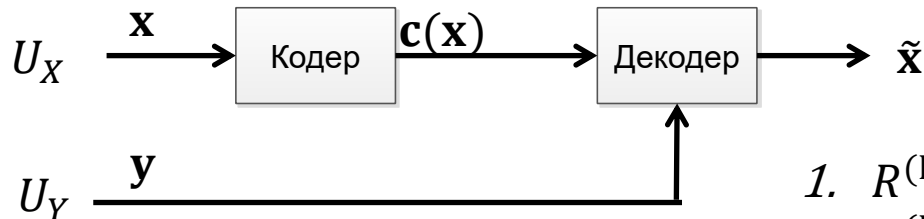
Кодирование зависимых источников с дополнительной информацией на декодере

А) Совместное кодирование и декодирование



$$R^{(A)}(d) = \min_{p(\mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}})} I(X; \tilde{X} | Y)$$

Б) Кодирование с дополнительной информацией на декодере (распределенное кодирование)

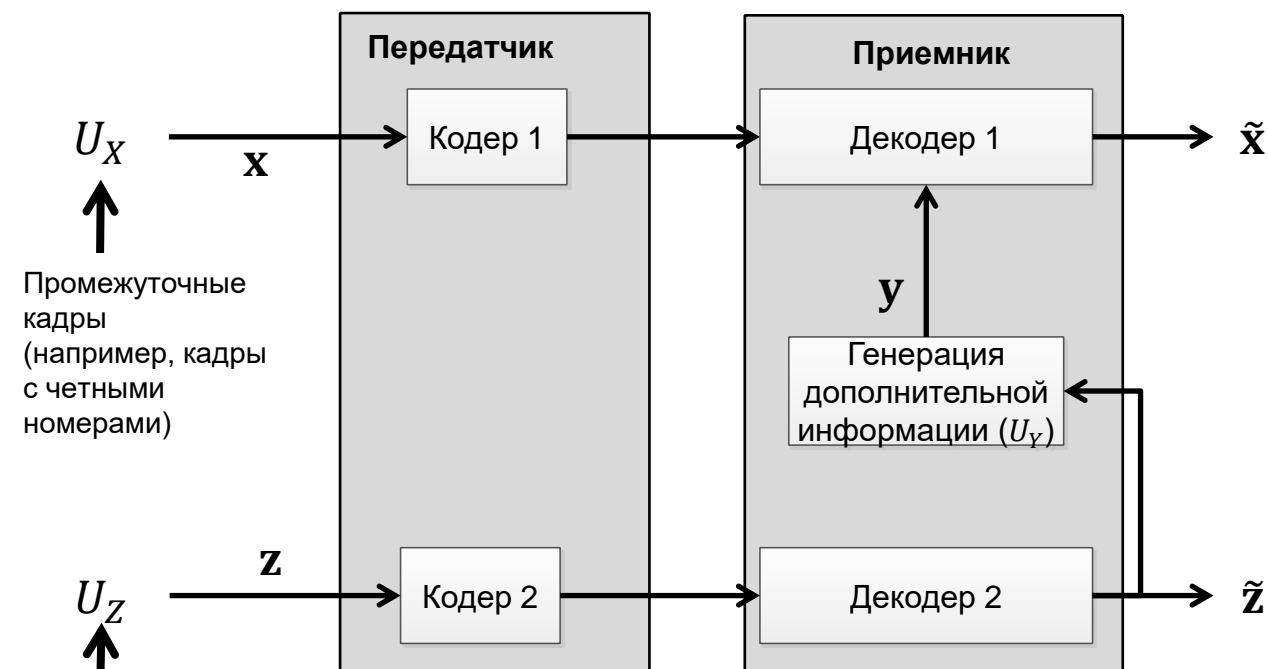


Граница Вайнера-Зива:

$$R^{(B)}(d) \geq R^{(A)}(d)$$

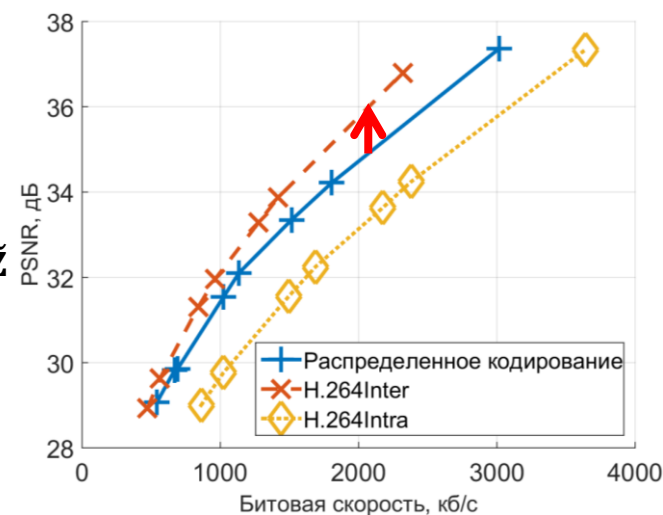
1. $R^{(B)}(0) = R^{(A)}(0) = H(X|Y)$
2. $R^{(B)}(d) = R^{(A)}(d)$, если $X \sim N(\mu_X, \sigma_X^2)$,
 $U \sim N(\mu_U, \sigma_U^2)$, $D(\mathbf{x}, \tilde{\mathbf{x}}) = (\mathbf{x} - \tilde{\mathbf{x}})^2$, $Y = X + U$

Применение распределенного кодирования для сжатия видео



Ошибка межкадрового
предсказания (correlation
noise):

$$\mathbf{n} = \mathbf{x} - \mathbf{y}.$$



Цель, объект и предмет исследования

- **Цель работы:** повышение степени сжатия без ухудшения качества восстановления видеоданных в кодеках, основанных на принципах кодирования зависимых источников с дополнительной информацией на декодере, за счет усовершенствования существующих и разработки новых способов обработки данных на стороне декодера.
- **Объектом** исследования является система сжатия видеоданных, основанная на принципах кодирования зависимых источников с дополнительной информацией на декодере.
- **Предмет** исследования составляет процесс восстановления промежуточных кадров на стороне декодера.

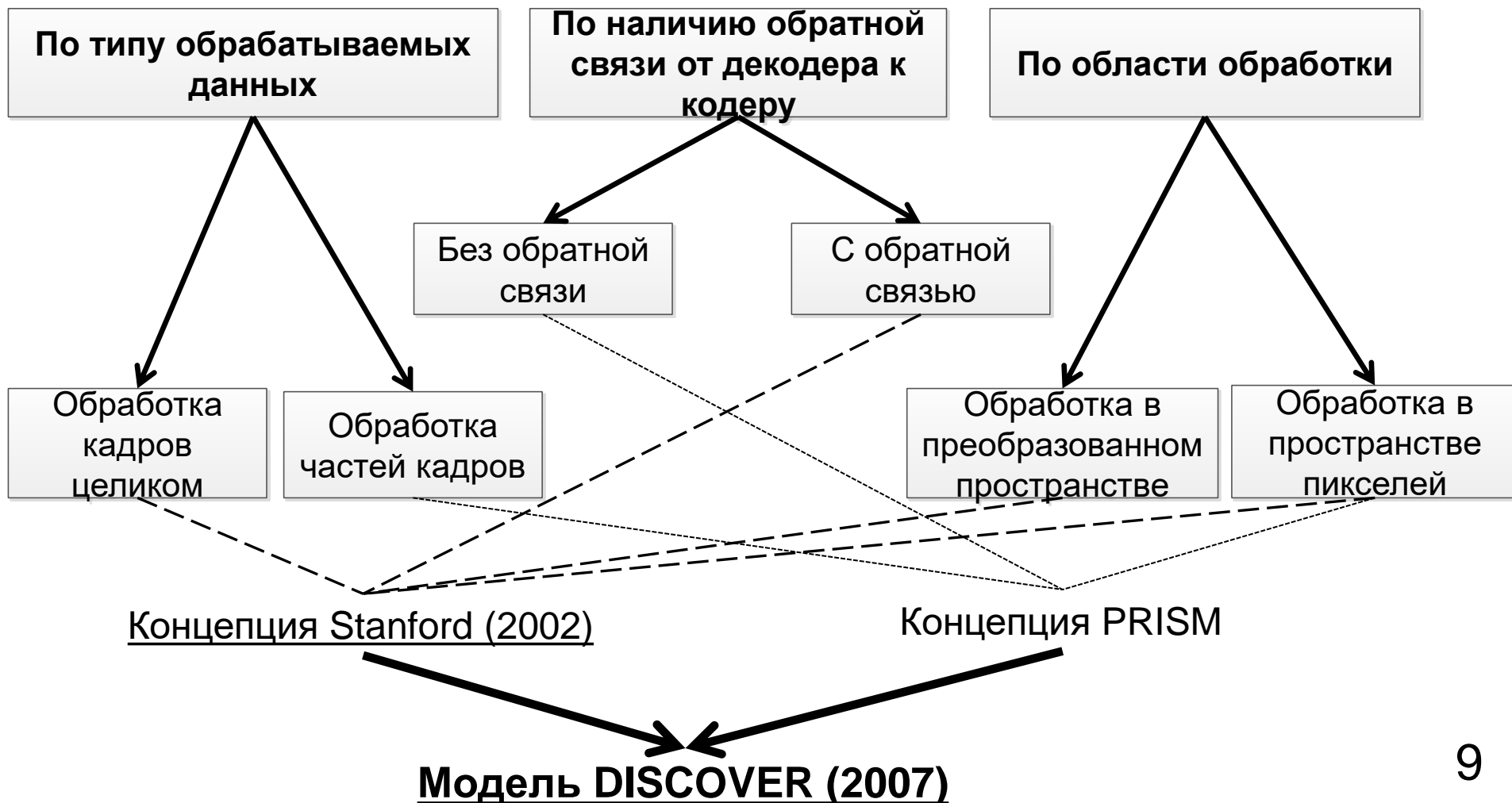
Задачи диссертационной работы

1. Исследовать типовые методы сжатия видеоданных, основанные на принципах распределенного кодирования источников.
2. Исследовать способы формирования дополнительной информации на стороне декодера.
3. Предложить новый алгоритм генерации дополнительной информации, учитывающий особенности входных данных.
4. Исследовать статистические характеристики ошибок межкадрового предсказания, возникающих при генерации дополнительной информации.
5. Разработать модель ошибок предсказания промежуточных кадров на стороне декодера.
6. Предложить алгоритм оценки параметров ошибок межкадрового предсказания.

Структура диссертационной работы

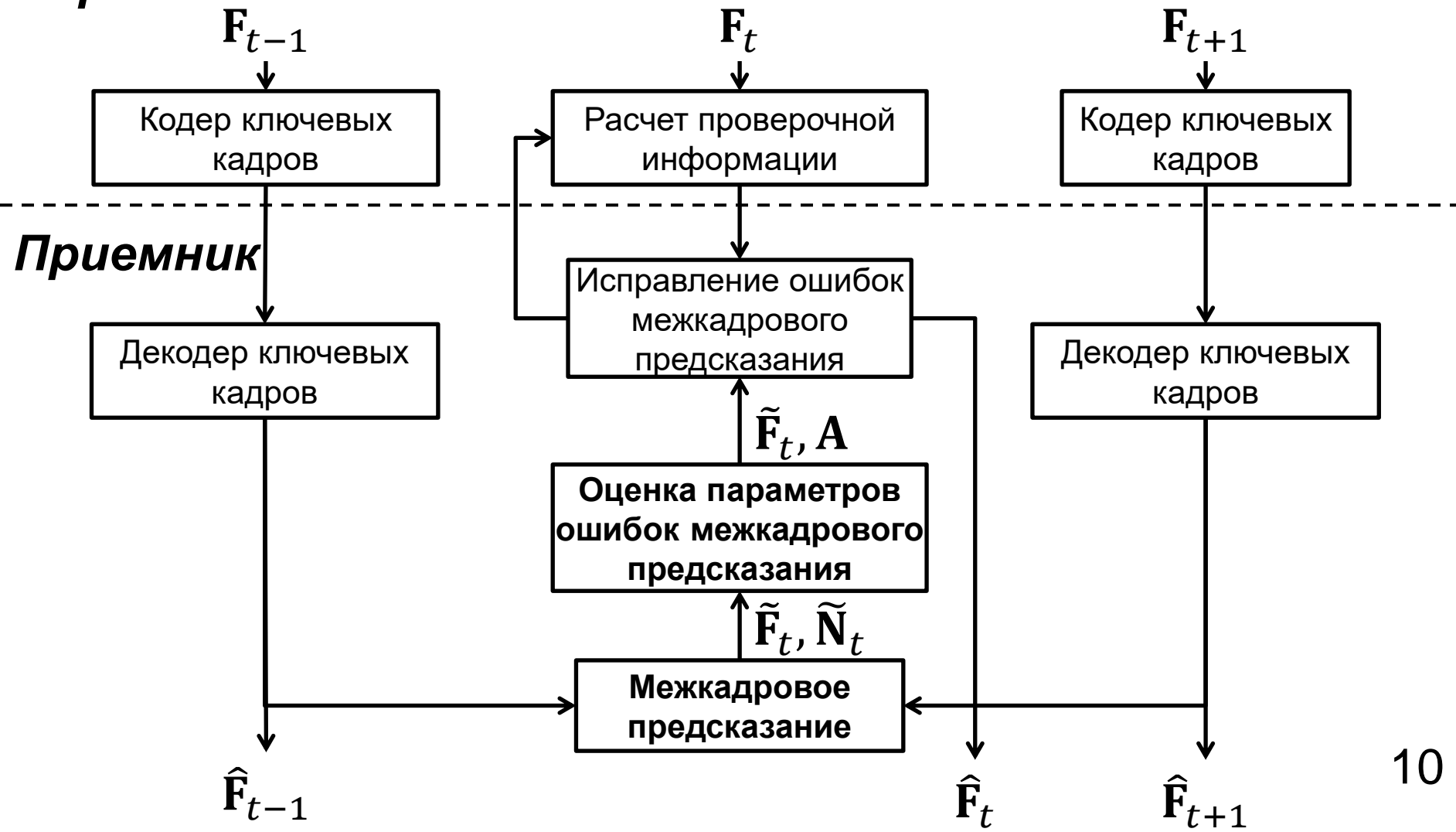
Раздел 1	<ul style="list-style-type: none">• Основные определения и обозначения.• Модель системы распределенного кодирования.• Основные факторы, влияющие на степень сжатия в системе распределенного кодирования.
Раздел 2	<ul style="list-style-type: none">• Модель истинного движения объектов в видеопоследовательности.• Анализ базового алгоритма генерации дополнительной информации.• Описание разработанного алгоритма генерации дополнительной информации.
Раздел 3	<ul style="list-style-type: none">• Анализ базового алгоритма оценки параметров ошибок межкадрового предсказания.• Модель виртуального канала.• Описание разработанного алгоритма оценки параметров ошибок межкадрового предсказания.
Раздел 4	<ul style="list-style-type: none">• Сравнительный анализ разработанных алгоритмов.• Метод сравнения алгоритмов генерации дополнительной информации.

Классификация методов распределенного кодирования источников видеоинформации

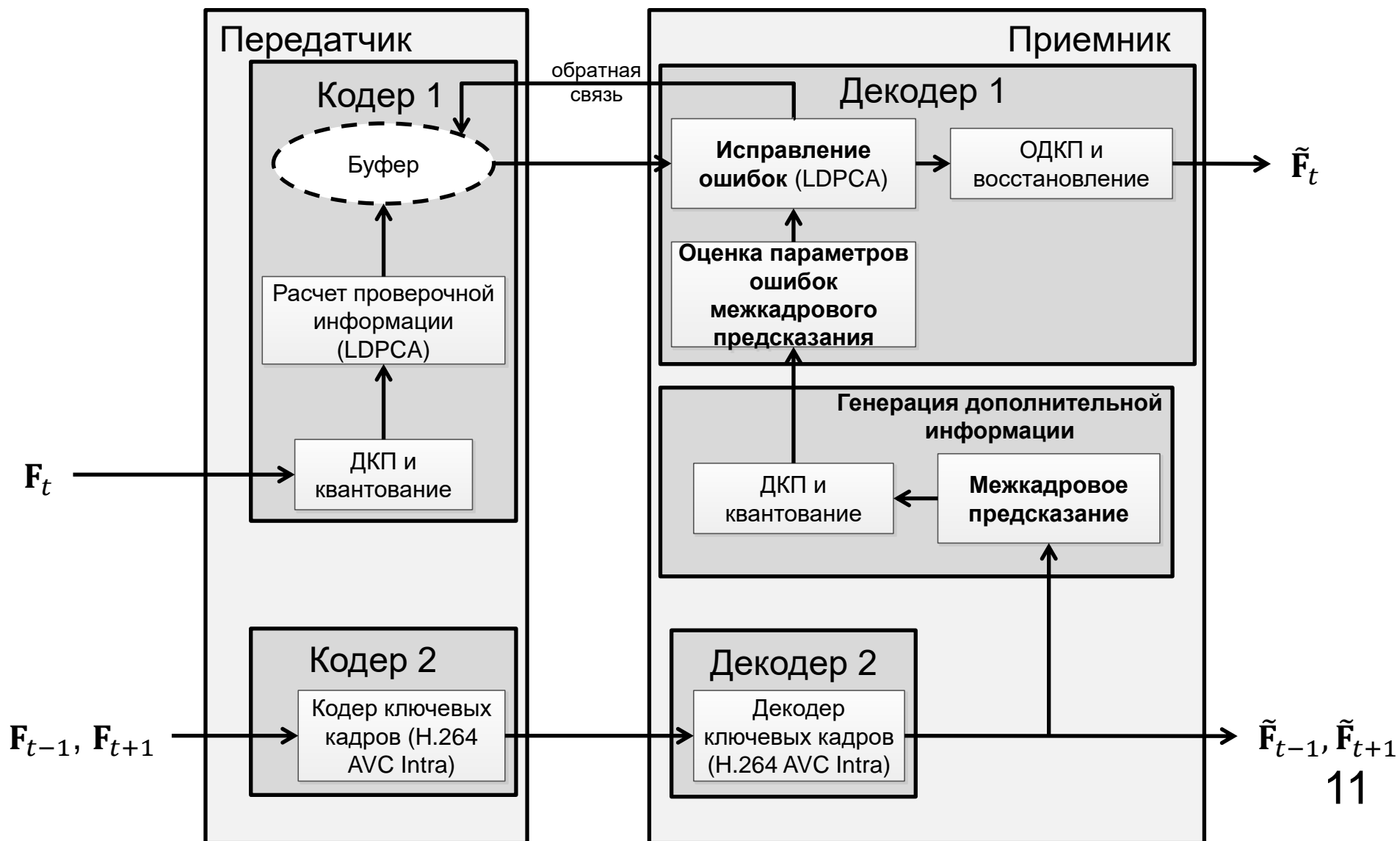


Процесс кодирования и декодирования кадров в DISCOVER

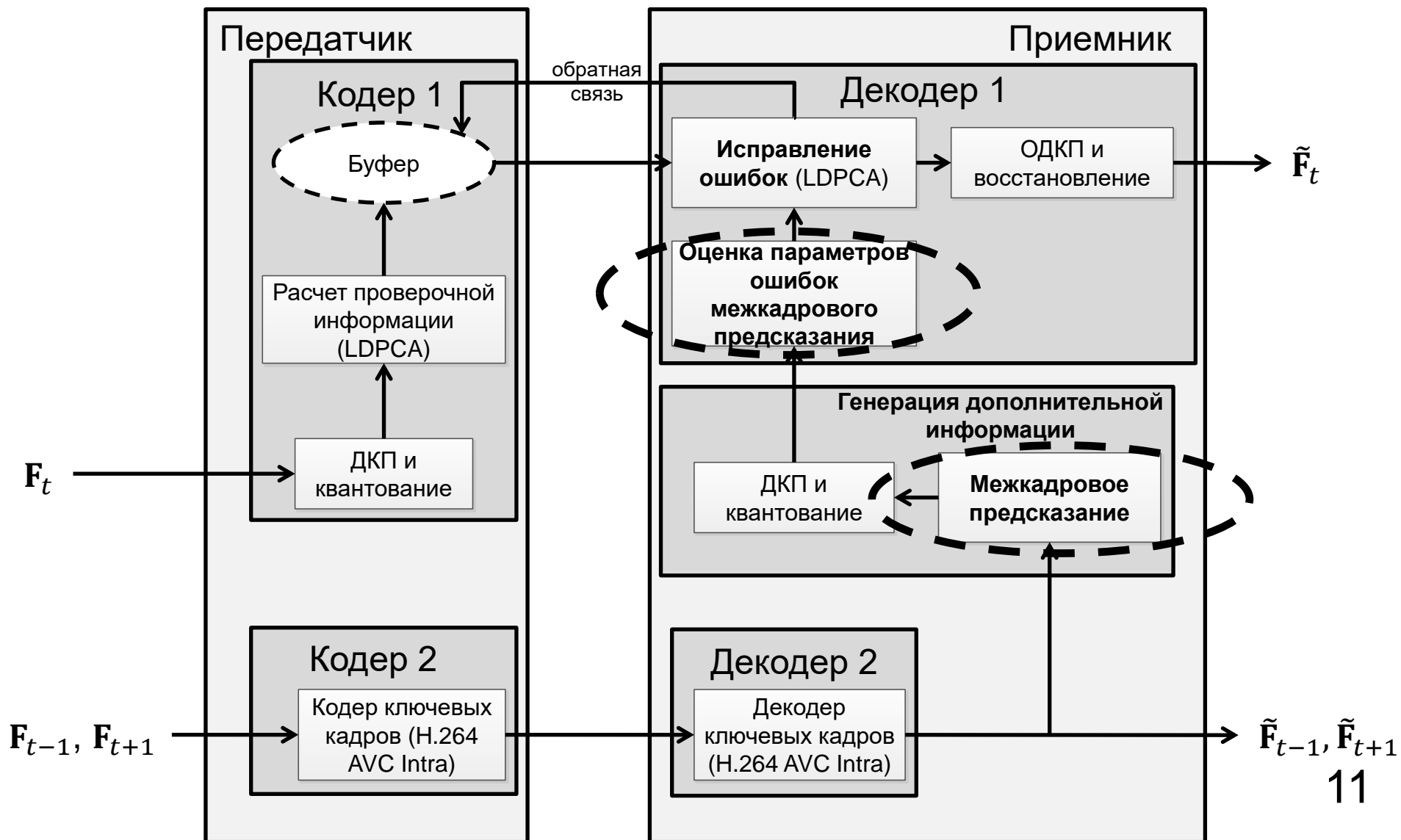
Передатчик



Основные модули базовой модели DISCOVER

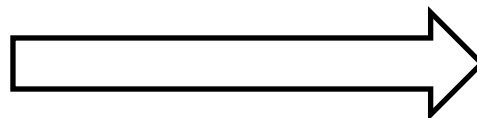
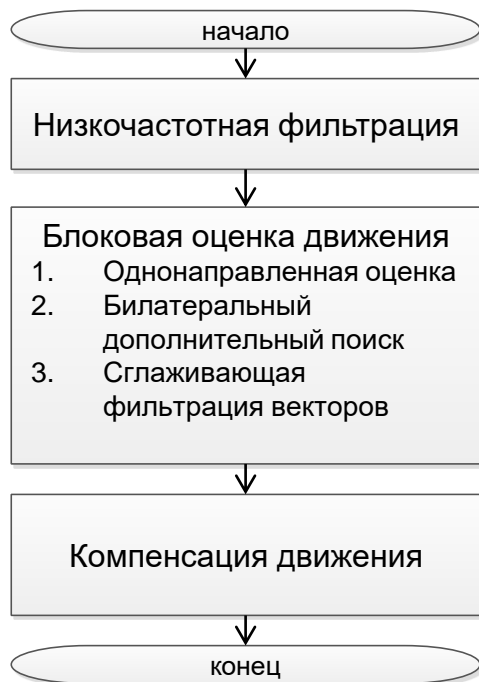


Основные модули базовой модели DISCOVER



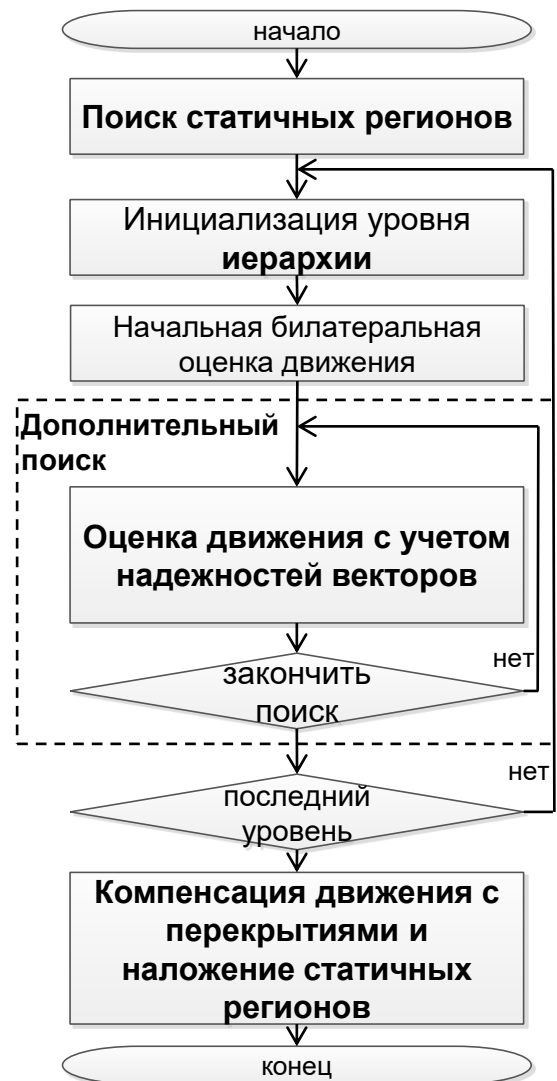
Разработка алгоритма межкадрового предсказания

Базовый алгоритм

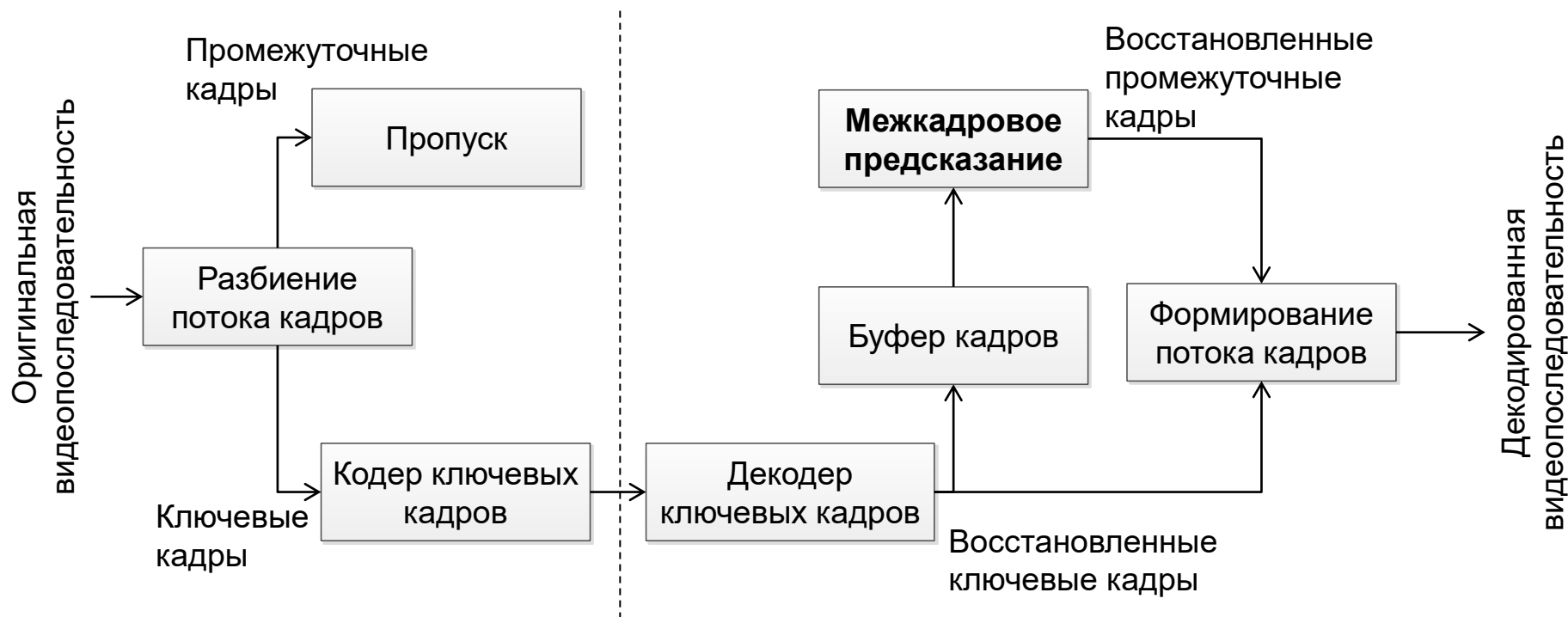


- Учет статических регионов
- Блоки переменного размера (иерархический подход)
- «Надежность» векторов
- Компенсация с перекрытиями

Разработанный алгоритм



Модель кодека для сравнения алгоритмов межкадрового предсказания

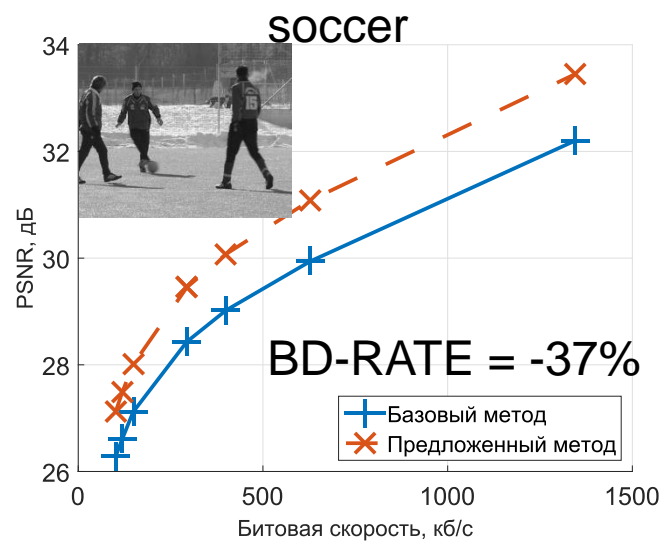
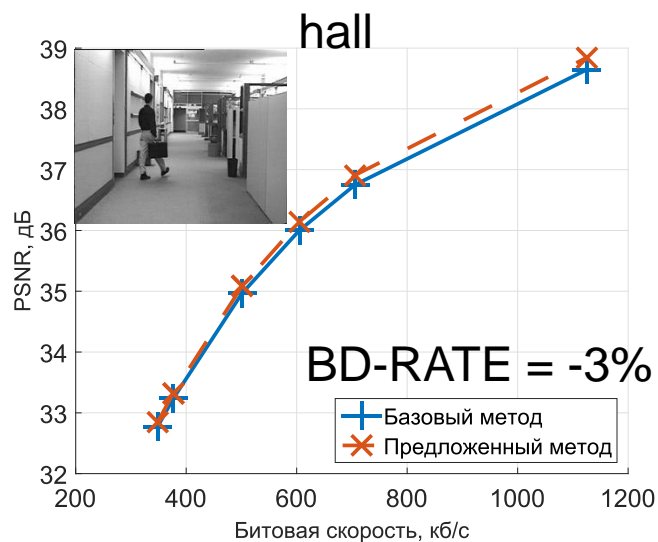
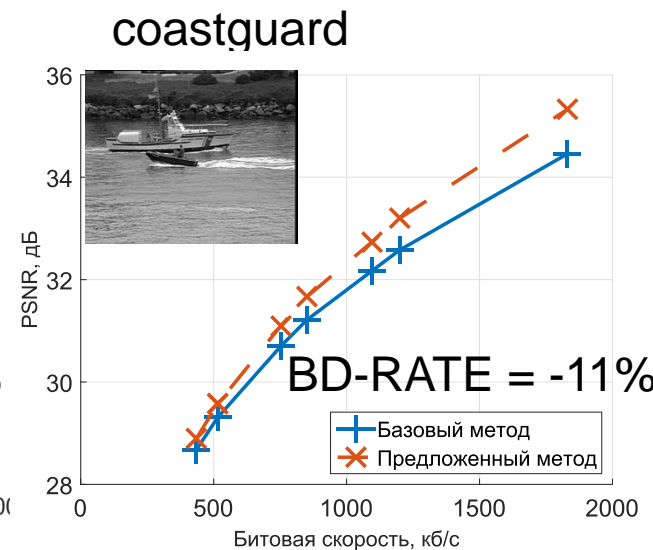
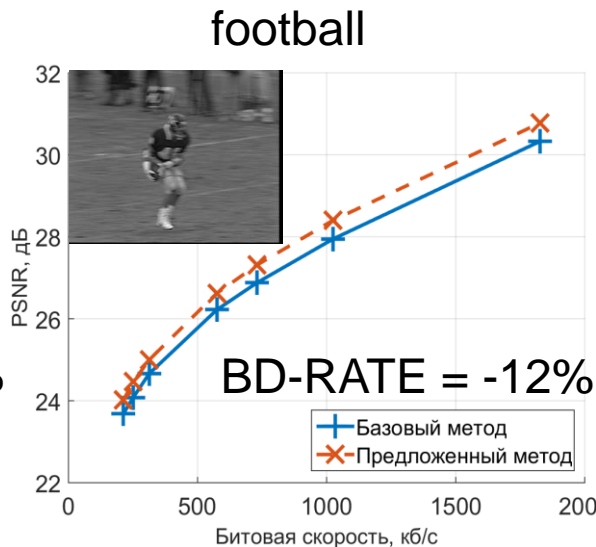
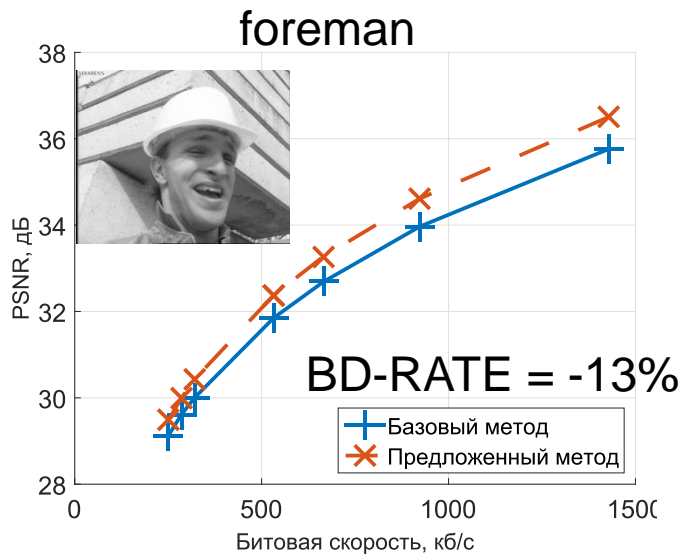


Характеристики тестового множества видеопоследовательностей

- Разрешение: 352x288
- Формат: YUV 4:0:0
- Кадровая скорость: 30 кадров в секунду

Последовательность	Число кадров	Экспертная оценка сложности движения
Coastguard	300	2
Football	260	4
Foreman	300	3
Hall	300	1
Soccer	300	4

Результаты сравнения алгоритмов межкадрового предсказания



Ошибки межкадрового предсказания в модели DISCOVER

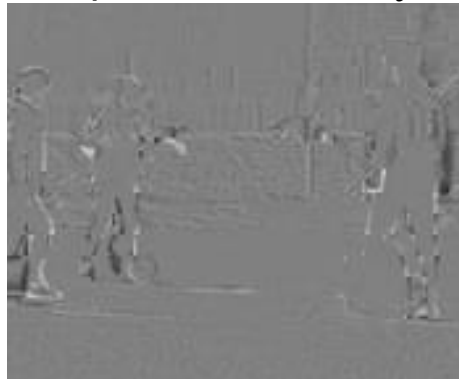
Промежуточный
кадр, F_t



Результат
межкадрового
предсказания, \tilde{F}_t



Аппроксимированные
ошибки межкадрового
предсказания, \tilde{N}_t



Восстановленный
кадр, \hat{F}_t



Ошибки межкадрового предсказания в модели DISCOVER

Кодер

Промежуточный кадр, F_t



ДКП 4x4 и
квантование

Расчет проверочной
информации
(LDPCA)

Буфер
кодера

$c^{(0,0)}$

$c^{(0,1)}$

...

$c^{(3,3)}$

Декодер

Результат
межкадрового
предсказания, \tilde{F}_t



ДКП 4x4 и
квантование

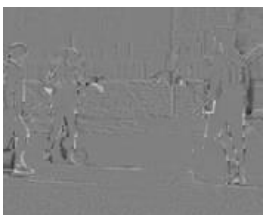
Исправление ошибок
(LDPCA)

ОДКП 4x4 и
восстановление

Восстановленный
кадр, \hat{F}_t



Аппроксимированные
ошибки межкадрового
предсказания, \tilde{N}_t

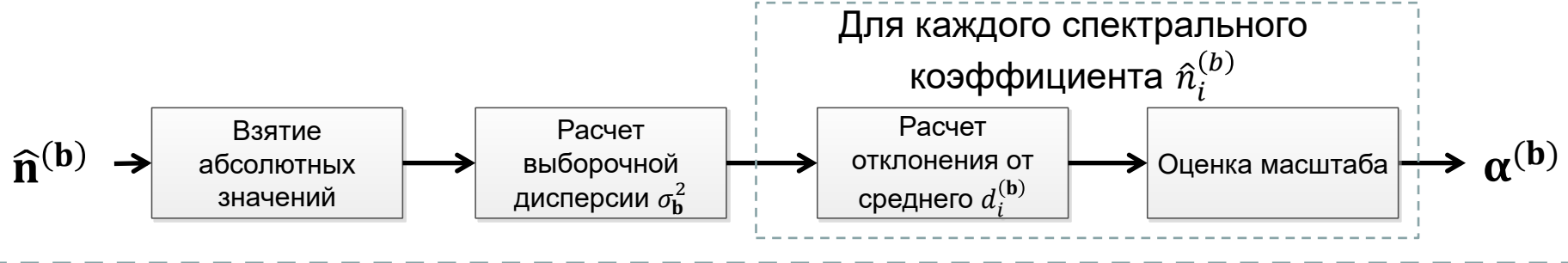


ДКП 4x4

Оценка параметров
ошибок межкадрового
предсказания

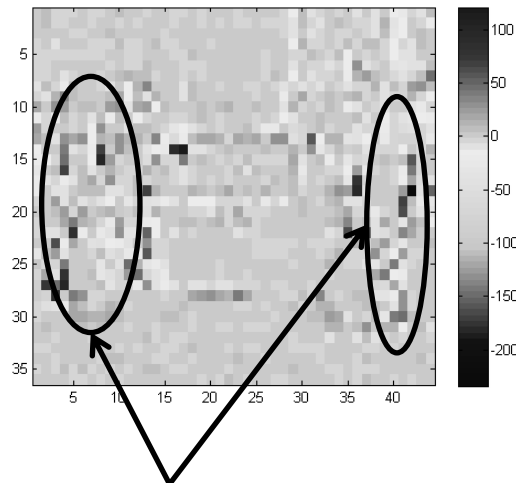
Базовый алгоритм оценки параметров ошибок межкадрового предсказания

$n_i^{(b)}$ — реализация случайной величины $N_i^{(b)} \sim \text{Lap}(\alpha_i^{(b)}, 0)$



Оценка масштаба:

$$\alpha_i^{(b)} = \begin{cases} \sqrt{\frac{2}{\sigma_b^2}}, & \text{если } (d_i^{(b)})^2 \leq \sigma_b^2 \\ \sqrt{\frac{2}{(d_i^{(b)})^2}}, & \text{иначе} \end{cases}$$

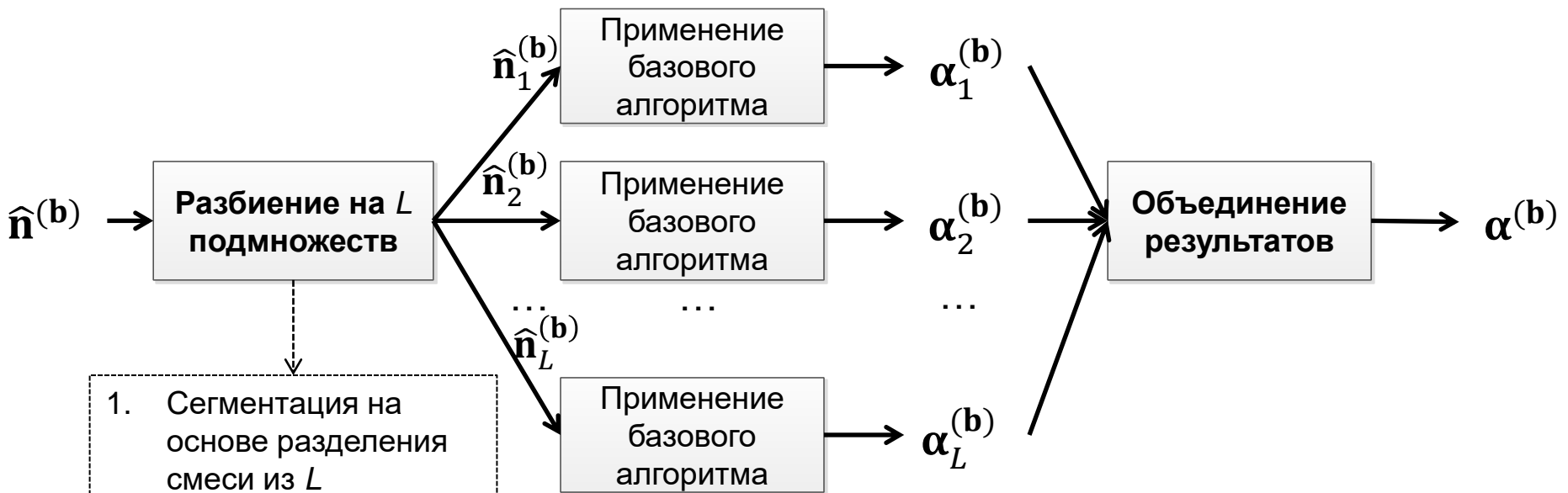


Идея модифицированного алгоритма.

1. Разбить множество ошибок на подмножества.
2. Проанализировать каждое подмножество независимо.

Группирование не учитывается в базовом алгоритме

Модифицированный алгоритм оценки параметров ошибок межкадрового предсказания



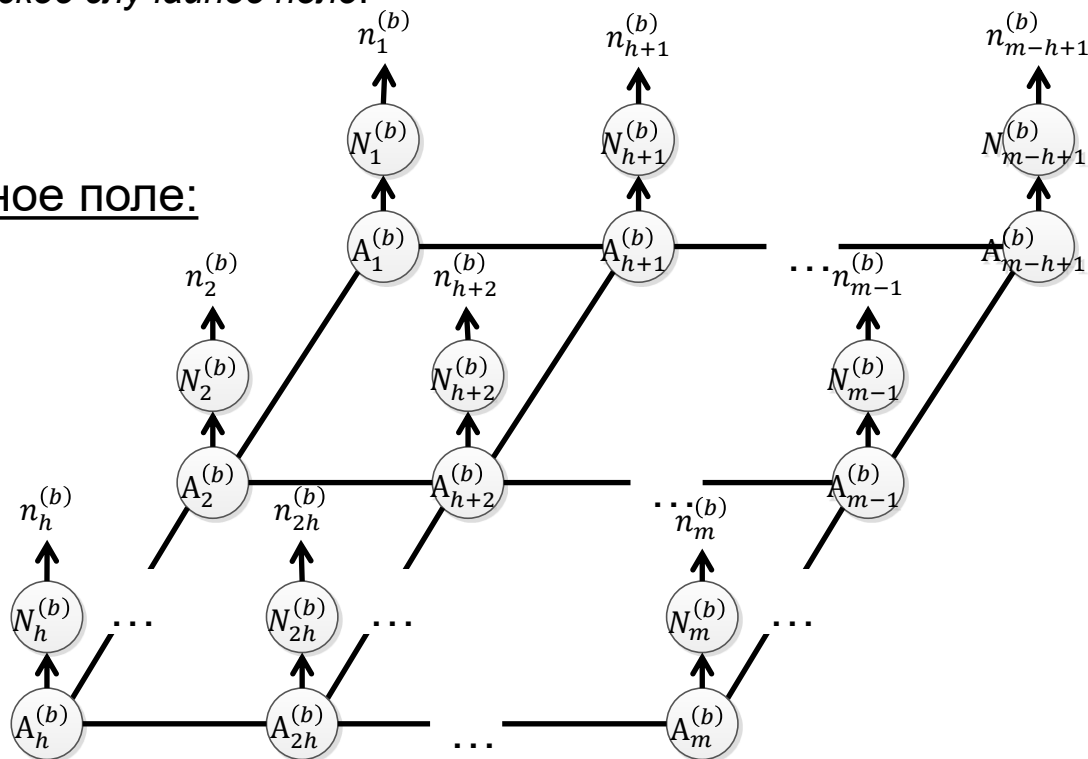
Выигрыш при использовании аппроксимированных значений ошибок

Последовательность	Выигрыш по критерию BD-Rate, %
Football	1 %
Soccer	0.6 %

Модель ошибок межкадрового предсказания

- **Определение 1.** Будем называть множество случайных величин $N^{(b)} = \{N_1^{(b)}, N_2^{(b)}, \dots, N_m^{(b)}\}$ *полем ошибок*.
- **Определение 2.** Будем называть множество случайных величин $A^{(b)} = \{A_1^{(b)}, A_2^{(b)}, \dots, A_m^{(b)}\}$, где $A_j^{(b)} \in \{1, 2, \dots, k^{(b)}\}$, *полем распределений*.
- **Допущение.** $A^{(b)}$ — Марковское случайное поле.

Скрытое Марковское случайное поле:



Оптимизационная задача для оценки параметров ошибок межкадрового предсказания

- Необходимо найти наиболее правдоподобный набор параметров с учетом Марковости поля:

$$\mathbf{a}^{(b)*} = \arg \max_{\mathbf{a} \in \{1,2,\dots,k^{(b)}\}^m} p(\mathbf{a}|\hat{\mathbf{n}}^{(b)}) = \arg \max_{\mathbf{a} \in \{1,2,\dots,k^{(b)}\}^m} p(\hat{\mathbf{n}}^{(b)}|\mathbf{a})p(\mathbf{a})$$

- Учитывая свойства условной независимости между случайными величинами:

$$\begin{aligned} \mathbf{a}^{(b)*} &= \arg \max_{\mathbf{a} \in \{1,2,\dots,k^{(b)}\}^m} p(\mathbf{a}) \prod_{i=1}^m p(\hat{n}_i^{(b)}|a_i) \\ &= \arg \max_{\mathbf{a} \in \{1,2,\dots,k^{(b)}\}^m} \prod_{i=1}^m p(\hat{n}_i^{(b)}|a_i) e^{-\frac{1}{t} \sum_{c \in C} \psi_c(\mathbf{a})}, \end{aligned}$$

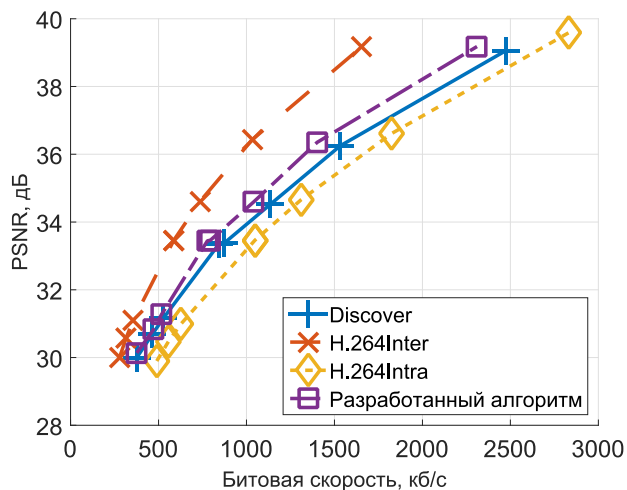
где $\psi_c(\mathbf{a})$ — ограничение Марковского поля

Сравнение с существующими кодеками

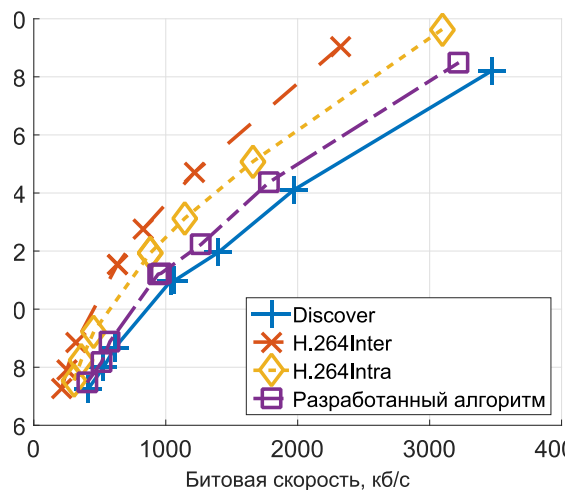
- Кодеки с низкой сложностью
 - H.264 Intra
 - DISCOVER
 - Разработанный кодек, включающий все предложенные алгоритмы
- Кодек с высокой сложностью
 - H.264 Inter

Результаты сравнения с существующими кодеками

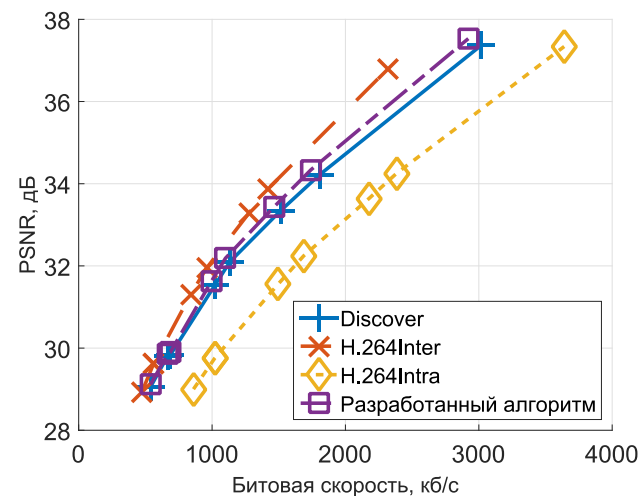
foreman



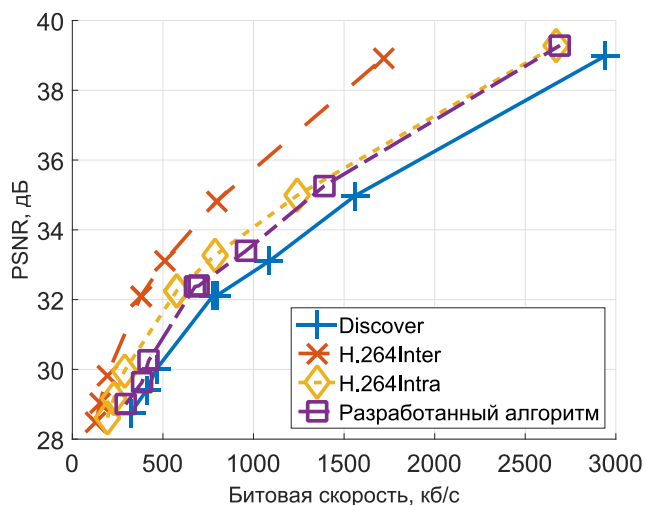
football



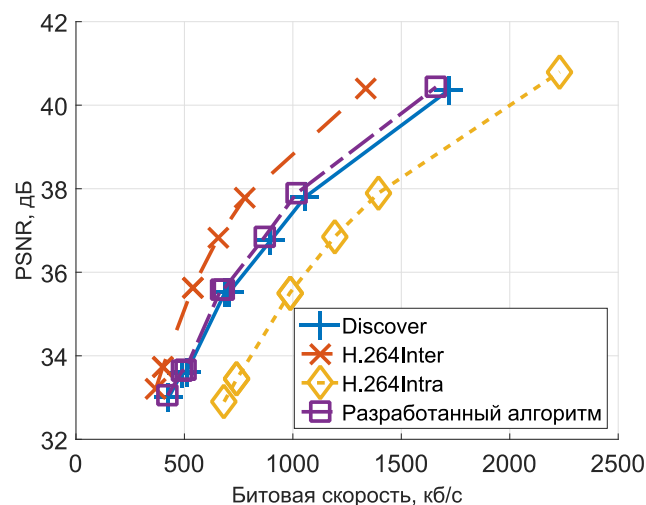
coastguard



soccer



hall



Сравнение с кодеком DISCOVER

Последовательность	Экспертная оценка сложности движения	BD-Rate, %
Coastguard	2	4,6
Football	4	12,4
Foreman	3	8,7
Hall	1	4,3
Soccer	4	15,8

Рекомендации по дальнейшему улучшению распределенного кодирования

- Применение аналитических моделей ошибок межкадрового предсказания для оценки степени сжатия при разработке алгоритмов распределенного кодирования;
- Оценка параметров ошибок межкадрового предсказания:
 - анализ распределений, отличных от распределения Лапласа;
 - учет шума квантования;
 - расчет аппроксимации ошибок межкадрового предсказания на стороне декодера.

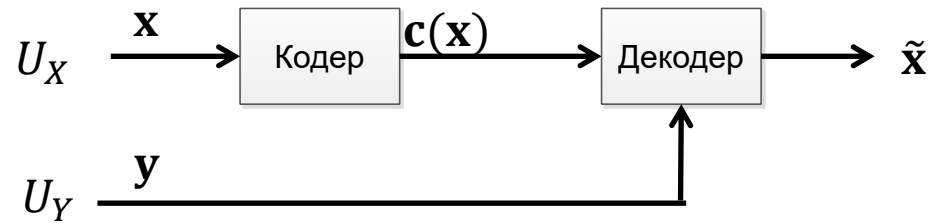
Внедрение результатов

- ЗАО «Интел А/О» в проекте «Разработка цепочки фильтров постобработки видеоданных»:
 - алгоритм поиска истинного движения в видеопоследовательностях;
 - алгоритм определения статичных регионов в видеопоследовательностях;
 - иерархический алгоритм временной интерполяции кадров.
- В учебном процессе кафедры инфокоммуникационных систем ГУАП:
 - курс «Мультимедиа технологии»;
 - курс «Цифровая обработка изображений».

Положения, выносимые на защиту

1. Алгоритм межкадрового предсказания для кодеков видеоинформации, основанных на принципах кодирования зависимых источников с дополнительной информацией на декодере, позволяющий уменьшить по сравнению с существующими алгоритмами число ошибок предсказания за счет использования временной интерполяции с учетом истинного движения объектов.
2. Модель кодека без обратной связи, позволяющая производить сравнение алгоритмов межкадрового предсказания в системах кодирования зависимых источников видеоинформации с дополнительной информацией на декодере.
3. Модифицированный алгоритм оценки параметров ошибок межкадрового предсказания в спектральной области, который за счет учета неоднородности ошибок в спектральных коэффициентах позволяет уменьшить битовые затраты на восстановление промежуточного кадра.
4. Модель ошибок межкадрового предсказания, основанная на Марковских случайных полях, которая позволяет учитывать пространственную зависимость между ошибками в спектральных коэффициентах.

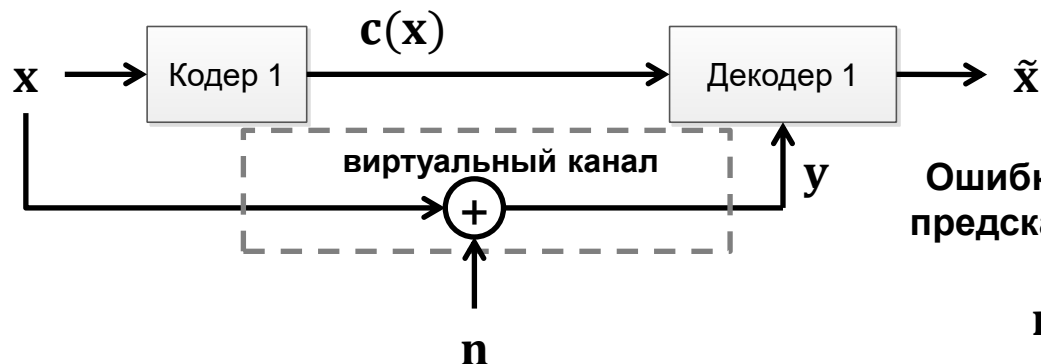
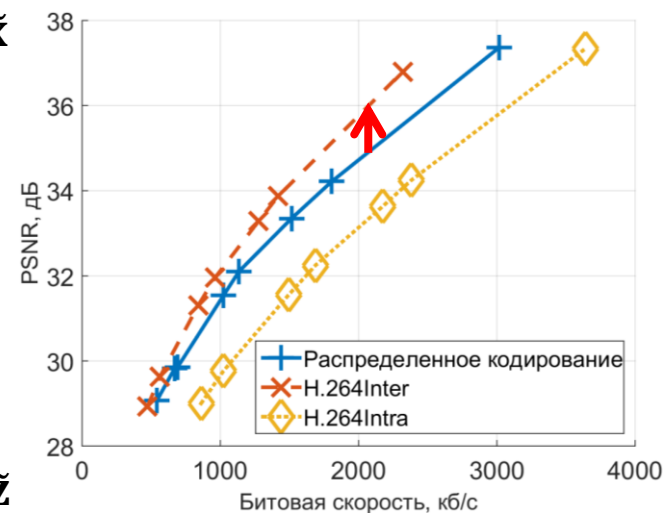
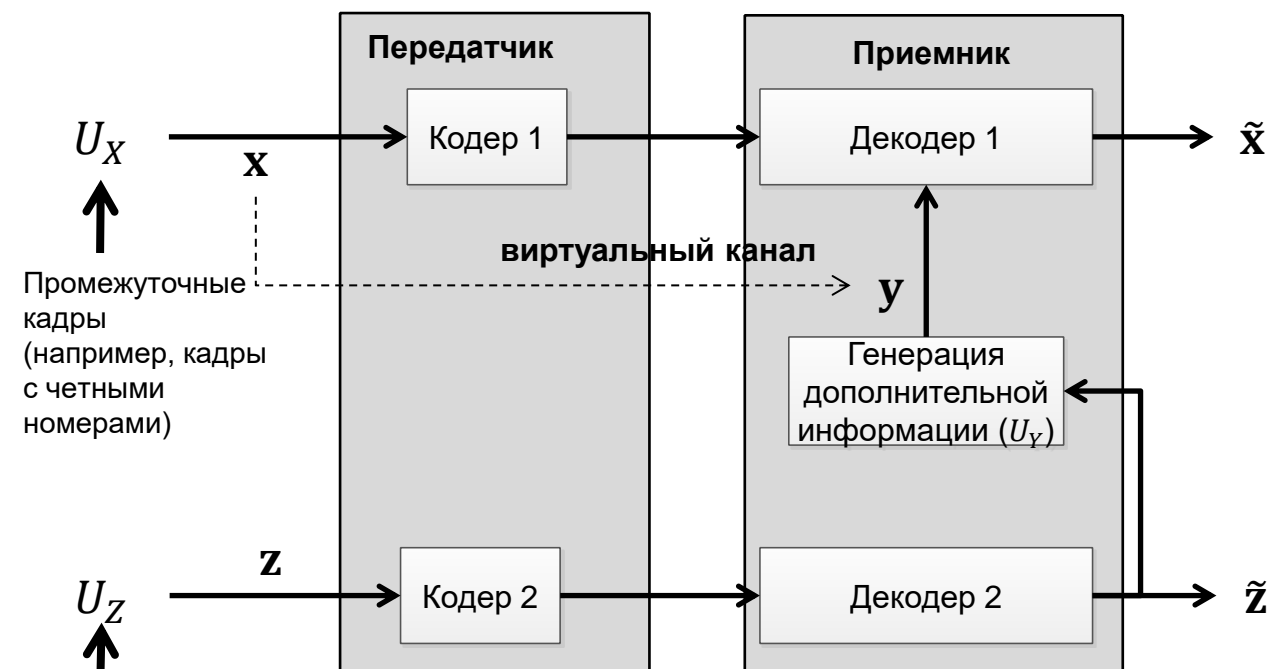
Пример распределенного кодирования



- Пусть $\mathbf{x}, \mathbf{y} \in \{0,1\}^3$ и $\mathbf{x} \oplus \mathbf{y} \leq \mathbf{1}$
- Для примера: $\mathbf{y} = (100)$, $\mathbf{x} = (101)$
- $\mathbf{H} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$
- Кодер: $\mathbf{c}(\mathbf{x}) = \mathbf{xH}^T = (10)$
- Декодер:
 - $(100) \oplus (010) = 2$
 - $(100) \oplus (101) = 1$

Синдром	Смежный класс
(00)	{(000), (111)}
(01)	{(001), (110)}
(10)	{(010), (101)}
(11)	{(100), (011)}

Применение распределенного кодирования для сжатия видео



Ошибка межкадрового предсказания (correlation noise):

$$n = x - y.$$

Ошибки межкадрового предсказания в модели DISCOVER

Кодер

Промежуточный кадр, F_t

ДКП 4x4 и квантование

$q^{(0,0)}$
 $q^{(0,1)}$
...
 $q^{(3,3)}$

Расчет проверочной информации (LDPCA)

Буфер кодера

$c^{(0,0)}$
 $c^{(0,1)}$
...
 $c^{(3,3)}$

Виртуальный канал

Декодер

Результат межкадрового предсказания, \tilde{F}_t

ДКП 4x4 и квантование

$\hat{q}^{(0,0)}$
 $\hat{q}^{(0,1)}$
...
 $\hat{q}^{(3,3)}$

Исправление ошибок (LDPCA)

$\tilde{q}^{(0,0)}$
 $\tilde{q}^{(0,1)}$
...
 $\tilde{q}^{(3,3)}$

ОДКП 4x4 и восстановление

Оценка ошибок межкадрового предсказания, \tilde{N}_t

ДКП 4x4

$\hat{n}^{(0,0)}$
 $\hat{n}^{(0,1)}$
...
 $\hat{n}^{(3,3)}$

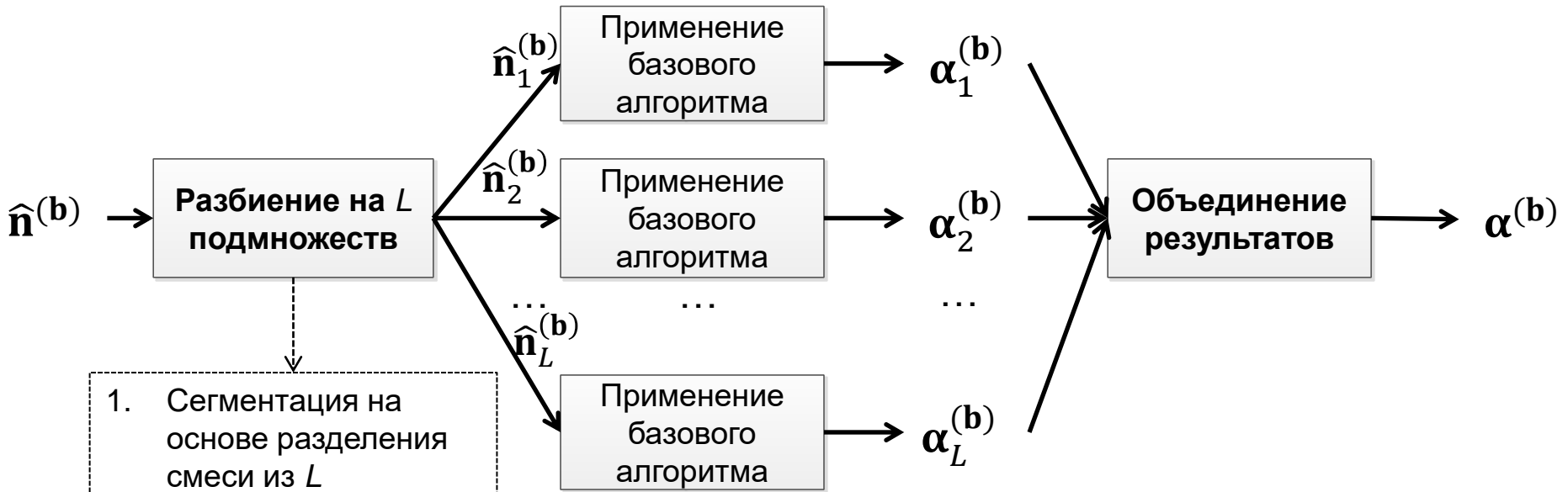
Оценка параметров ошибок межкадрового предсказания

$\hat{\alpha}^{(0,0)}$ $\hat{\alpha}^{(0,1)}$... $\hat{\alpha}^{(3,3)}$

Декодирование коэффициентов каждого подмножества делается независимо

Восстановленный кадр, \hat{F}_t

Модифицированный алгоритм оценки параметров ошибок межкадрового предсказания



Выигрыш при использовании реальных значений ошибок

Последовательность	Выигрыш по критерию BD-Rate, %
Football	3.8 %
Soccer	2.2 %

Раскрытие положений, выносимых на защиту, в презентации

- Положение 1: слайд 12
- Положение 2: слайд 18-19
- Положение 3: слайд 20
- Положение 4: слайд 13

Критерий Бьёнтегаарда (BD-Rate)

1. Кривые «скорость-искажение» аппроксимируются полиномами третьей степени:

$$\log_2 R(D) = w_3 D^3 + w_2 D^2 + w_1 D + w_0$$

2. Разница между кривыми оценивается как

$$\Delta R = M \left[\frac{R_2(D) - R_1(D)}{R_1(D)} \right] \approx 10^{\frac{1}{D_h - D_l}} \int_{D_l}^{D_h} (\log_2 R_2(D) - \log_2 R_1(D)) dD,$$

где

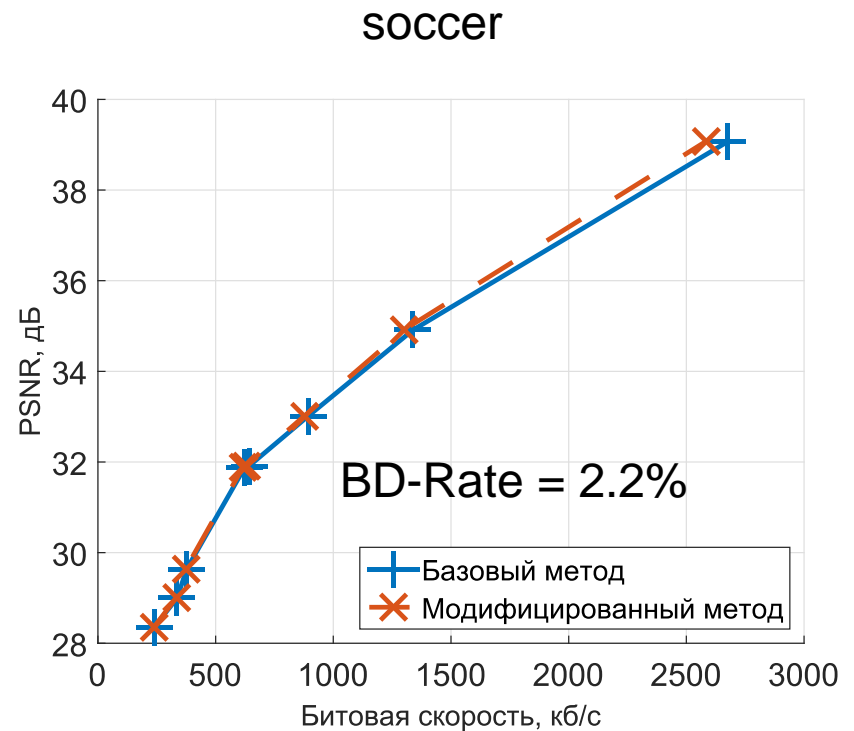
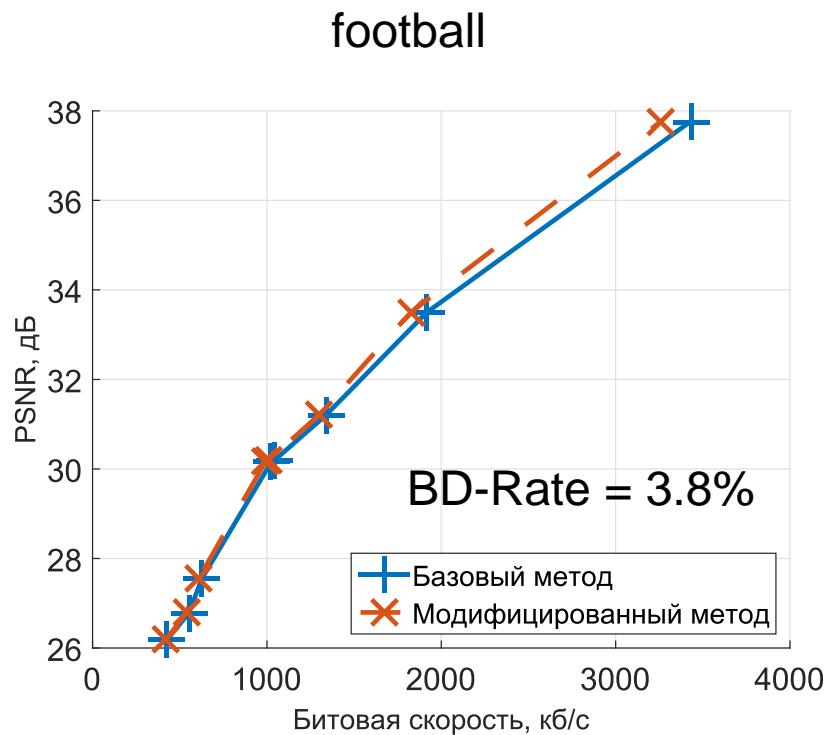
$$D_l = \max(\min(D_{1,1}, D_{1,2}, \dots, D_{1,N}), \min(D_{2,1}, D_{2,2}, \dots, D_{2,N})),$$

$$D_h = \min(\max(D_{1,1}, D_{1,2}, \dots, D_{1,N}), \max(D_{2,1}, D_{2,2}, \dots, D_{2,N})).$$

Основные результаты работы

1. Разработан алгоритм генерации дополнительной информации на стороне декодера, использующий метод временной интерполяции, основанный на процедуре поиска истинного движения объектов в видеопоследовательности.
2. Предложена модель виртуального канала, основанная на вероятностной порождающей модели ошибок, представляющей процесс возникновения искажений с помощью скрытого Марковского случайного поля (СМСП).
3. Разработан модифицированный алгоритм моделирования корреляционного шума, основанный на процедуре сегментации ошибок межкадрового предсказания с использованием ЕМ-подхода.
4. Предложен метод выполнения сравнительной оценки алгоритмов генерации дополнительной информации, учитывающий тот факт, что модули системы распределенного кодирования оказывают сильное влияние друг на друга.
5. Предложен метод построения верхней границы кривых «скорость-искажение», показывающей для заданной степени сжатия максимально возможное качество восстановления кадров при использовании билатеральной оценки движения в модуле генерации дополнительной информации.

Сравнение алгоритмов оценки параметров ошибок межкадрового предсказания (реальные значения ошибок)



Модель истинного движения

$$E(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}, \alpha) = E_{res}(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}) + \alpha E_{smooth}(\mathbf{V}),$$

где α – коэффициент регуляризации, E_{res} – энергия разностного кадра, для векторного поля \mathbf{V} , E_{smooth} – энергия векторного поля.

$$\mathbf{V} = \mathbf{argmin}_{\mathbf{V} \in \mathcal{V}} E(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}, \alpha)$$

Базовый алгоритм с учетом модели истинного движения.

1. Для заданного размера блока найти билатеральное векторное поле с минимальной ошибкой сопоставления блоков (минимизация E_{res}).
2. Сгладить векторное поле (уменьшение E_{smooth}).

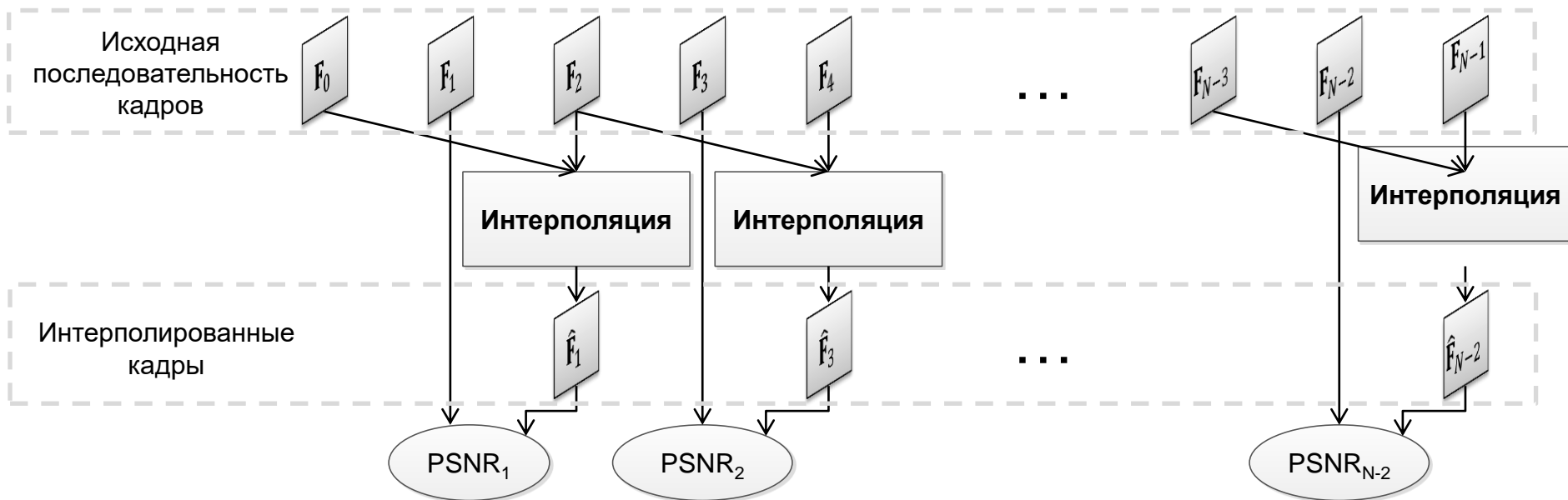
Основные недостатки базового алгоритма

- Использование блоков только одного фиксированного размера не позволяет осуществлять поиск с учетом глобального движения.
- Сглаживание поля не учитывает «надежность» векторов движения.

Основные влияющие на сжатие факторы

- **Точность генерации дополнительной информации:** чем меньше различий между дополнительной информацией декодера и исходными данными на стороне кодера, тем меньше проверочных бит необходимо затратить на их исправление.
- **Эффективность исправления ошибок** в дополнительной информации, на которую влияют:
 - **модуль оценки параметров виртуального канала:** надежности символов оказывают существенное влияние на эффективность исправления ошибок с использованием корректирующих кодов;
 - **модуль помехоустойчивого кодирования:** чем выше корректирующая способность кода, тем больше ошибок он позволяет исправить при фиксированной длине кода.

Результаты сравнения предложенного алгоритма межкадрового предсказания с базовым



Последовательность	Среднее значение PSNR, дБ	
	Базовый алгоритм	Разработанный алгоритм
COASTGUARD	32,12	34,66
FOREMAN	32,55	34,25
HALL	36,89	37,35
SOCCER	25,13	27,66

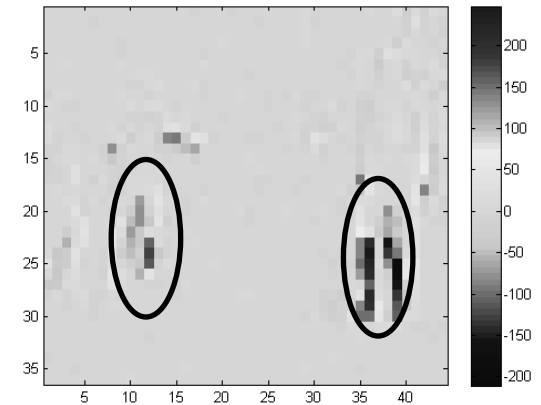
Базовые допущения при оценке параметров ошибок межкадрового предсказания

- **Допущение 1.** $n_i^{(b)}$ – случайная величина из некоторого вероятностного закона, плотность которого можно аппроксимировать как

$$f(x) = \sum_{i=1}^{k^{(b)}} w_i^{(b)} \text{Lap}(x|\alpha_i, 0).$$

- **Допущение 2.** Декодер может рассчитать аппроксимацию шума $\hat{\mathbf{n}}^{(b)} = (\hat{n}_1^{(b)}, \hat{n}_2^{(b)}, \dots, \hat{n}_m^{(b)})$, причем $\hat{n}_i^{(b)} = n_i^{(b)} + e_i^{(b)}$, где $e_i^{(b)}$ – случайное искажение.

Реальный шум (DC)



Аппроксимация шума (DC)

