

Исследование границ интенсивности видеопотока при FPV-управлении БПЛА в режиме предсказания кадров. Часть II: адаптивный контроль интенсивности видеопотока

А. А. Березкин, А. А. Ченский, Р. В. Киричек

Санкт-Петербургский университет телекоммуникаций им. проф. М. А. Бонч-Бруевича (СПбГУТ)

Аннотация: В настоящее время беспилотные системы активно используются во многих сферах народного хозяйства, в том числе в режиме управления от первого лица. Для обеспечения управления от первого лица необходима передача видеопотока от беспилотной системы к станции внешнего пилота. Тем не менее, при передаче видеопотока через гибридные орбитально-наземные сети связи возникает большая задержка и возможна потеря кадров видеопотока. Это приводит к снижению интенсивности видеопотока и увеличению задержки отображения видеопотока на терминале оператора. В предыдущей части исследования рассматривался метод повышения интенсивности видеопотока, были построены его математические модели. В связанных работах исследовалась возможность уменьшения времени рассинхронизации управления от первого лица, а также представлена структура системы информационного обмена для управления беспилотными системами. Во второй части настоящего исследования обобщаются полученные в ранних работах авторов результаты и рассматривается система адаптивного контроля интенсивности видеопотока, которая работает в двух режимах: повышения интенсивности и уменьшения задержки при передаче видеопотока через космический сегмент гибридной сети связи.

Ключевые слова: нейронная сеть, вариационный автокодировщик, управление от первого лица, FPV, беспилотная авиационная система, предсказание видеопотока, DMVFN, интенсивность видеопотока.

Для цитирования: Березкин А. А., Ченский А. А., Киричек Р. В. Исследование границ интенсивности видеопотока при FPV-управлении БПЛА в режиме предсказания кадров. Часть II: адаптивный контроль интенсивности видеопотока // Вестник СибГУТИ. 2025. Т. 19, № 3. С. 139–164. <https://doi.org/10.55648/1998-6920-2025-19-3-139-164>.



Контент доступен под лицензией
Creative Commons Attribution 4.0
License

© Березкин А. А., Ченский А. А.,
Киричек Р. В., 2025

Статья поступила в редакцию 13.03.2025;
переработанный вариант – 28.04.2025;
принята к публикации 28.04.2025.

1. Введение

На качество управления беспилотными воздушными судами (БВС) от первого лица в первую очередь влияет интенсивность кадров видеопотока на стороне оператора и рассинхронизация поступающего видео и команд FPV-управления (first person view control). При низкой интенсивности видеопотока возникают большие паузы между кадрами видеопотока, поступающими на станцию внешнего пилота (СВП). При интенсивности видеопотока FPS кадров в секунду период между поступлением кадров видеопотока на СВП приблизительно равен

$$T = \frac{1000}{FPS} \text{ (мс)}. \quad (1)$$

Пусть во время FPV-управления возникает некоторое нештатное событие C_1 и время, через которое отсутствие реакции на него приведёт к недостижению цели функционирования БВС, равно TC_1 . Пусть событие возникло через Δtc_1 после отправки на СВП очередного кадра x_i во момент времени t_i . При этом существует задержка видеопотока T_3^B и задержка команд управления T_3^Y .

Если $TC_1 - \Delta tc_1 < T$, то своевременная реакция невозможна. Кадр x_i не отражает событие C_1 . Следовательно, для формирования реакции внешнего пилота на событие C_1 требуется поступление кадра x_{i+1} , но при достаточно низкой интенсивности видеопотока кадр не поступит в срок для предотвращения срыва полётного задания.

Более того, условие $TC_1 - \Delta tc_1 \geq T$ не является достаточным для правильной реакции на событие C_1 . Следующий кадр видеопотока x_{i+1} , отражающий событие C_1 в достаточной степени для корректной интерпретации внешним пилотом (ВП), будет отправлен с БВС в момент времени $t_{i+1} = t_i + T$ и получен на СВП в момент времени

$$\hat{t}_{i+1} = t_{i+1} + T_3^B = t_i + T + T_3^B,$$

где T_3^B – задержка передачи видеопотока.

Команды управления в ответ на полётную обстановку x_{i+1} и событие C_1 поступят на БВС в

$$\hat{t}_{i+1} + T_3^Y = t_{i+1} + T_3^B = t_i + T + T_3^B + T_{\text{реакт}} + T_3^Y,$$

где $T_{\text{реакт}}$ – время реакции ВП, T_3^Y – задержка передачи команд управления.

Для достижения цели функционирования БВС должно выполняться условие

$$t_i + T + T_3^B + T_{\text{реакт}} + T_3^Y < t_i + \Delta tc_1 + TC_1,$$

$$T + T_3^B + T_{\text{реакт}} + T_3^Y < \Delta tc_1 + TC_1.$$

Следовательно, интервал времени реакции ВП равен

$$T_{\text{реакт}} \in [0; \Delta tc_1 + TC_1 - T - T_3^B - T_3^Y] \text{ (мс)} \quad (2)$$

При условии $\Delta tc_1 + TC_1 \leq T + T_3^B + T_3^Y$ интервал вырождается и реакция становится невозможной. Таким образом, чем выше интенсивность видеопотока, тем больше у ВП времени на принятие решения, адекватного полетной обстановке. Однако, увеличение интенсивности видеопотока приводит только к уменьшению T (1). Дальнейшее увеличение интервала времени реакции ВП достигается за счет уменьшения рассинхронизации FPV-управления [1]:

$$T_p = T_3^B + T_3^Y.$$

Выражение (2) сводится в этом случае к

$$T_{\text{реакт}} \in [0; \Delta t c_1 + TC_1 - T - T_p] \text{ (мс)}.$$

К уменьшению интенсивности видеопотока приводят потери пакетов в сети связи, а к увеличению рассинхронизации FPV-управления – сетевая задержка. При этом для повышения вероятности достижения цели функционирования БВС необходимо как повышение интенсивности видеопотока, так и уменьшение рассинхронизации FPV-управления. Настоящая работа посвящена устранению данного противоречия. В ней представлены модели и методы для повышения интенсивности видеопотока и уменьшения рассинхронизации FPV-управления.

В первой части работы [2] был предложен и исследован метод увеличения интенсивности видеопотока FPV-управления БВС, разработана регрессионная модель времени предсказания кадров видеопотока с помощью нейросетевой модели DMVFN (Dynamic Multi-scale Voxel Flow Network – динамическая многомасштабная нейронная сеть потока вокселей) [3] в зависимости от количества предсказываемых пикселей, а также представлена математическая модель для оценки возможности увеличения интенсивности видеопотока методом предсказания. Был достигнут важный результат: экспериментально и формально была доказана возможность безусловного увеличения интенсивности видеопотока предложенным методом при выполнении на видеокарте Nvidia A100 80Gb [4, 5] и предсказании кадров разрешения 512x512. Тем не менее, предложенный метод неизбежно приводит к увеличению рассинхронизации FPV-управления на время предсказания кадров T_p .

В [1] предложен метод уменьшения времени рассинхронизации FPV-управления путём предсказания будущих кадров и дана оценка применимости данного метода. В отдельных случаях при использовании метода возможно «упреждающее FPV-управление» – поступление команд управления на БВС до того, как будет отправлен кадр с полётной обстановкой, на основе которой были сформированы данные команды. Существует высокая корреляция между динамикой видеопотока и качеством восстанавливаемых кадров. Так на высокодинамичном видеопотоке на динамичных участках удалось предсказывать кадры на 40 мс вперёд, а на статичных участках – на 80–120 мс вперёд с допустимым качеством. На более статичных видеопотоках следует ожидать лучших результатов. Но при передаче кадров видеопотока через спутниковые системы связи метод имеет ограниченное применение. На низкой околоземной орбите (НОО) и отчасти на средней околоземной орбите (СОО) сетевые задержки (в основном за счёт задержки распространения, определяемого скоростью света в вакууме) составляют приблизительно 50 и 180–360 мс [6]. Данные задержки могут в существенной степени компенсироваться методом. Но на геостационарной земной орбите (ГОО) и высокой околоземной орбите (ВОО) сетевые задержки уже равны 600 мс и выше [6], и при достаточно высокой динамичности видеопотока метод позволяет компенсировать лишь малую их часть.

Для предсказания кадров в исследовании [1] используется два кадра видеопотока. Если ни одно из них не включает признаковые структуры некоторого события C_i , которое должно быть отображено в предсказываемом кадре $\hat{x}_i$, то данное событие не будет в него включено. Это свойство уменьшает значимость данного метода, так как ограничивает возможность предсказывать события C_i , угрожающие сорвать выполнение полётного задания.

В настоящей работе на базе данных методов разрабатываются блоки контроля интенсивности видеопотока (КИВ) и анализатора каналов информационного обмена (КИО). Данные блоки включают в себя функциональность повышения интенсивности видеопотока и уменьшения рассинхронизации FPV-управления.

2. Система информационного обмена для управления БС

В [7] была представлена структура системы информационного обмена беспилотных систем (БС) в гибридных орбитально-наземных сетях связи (ГОНСС) (рисунок 1) при использовании нейросетевых методов кодирования видеопотока и агрегации команд управления. ГОНСС представляют собой сети связи, интегрирующих наземные и космические сегменты. Согласно Стратегии развития отрасли связи Российской Федерации на период до 2035 года [8], появится возможность FPV-управления на всей территории Российской Федерации с использованием ГОНСС.

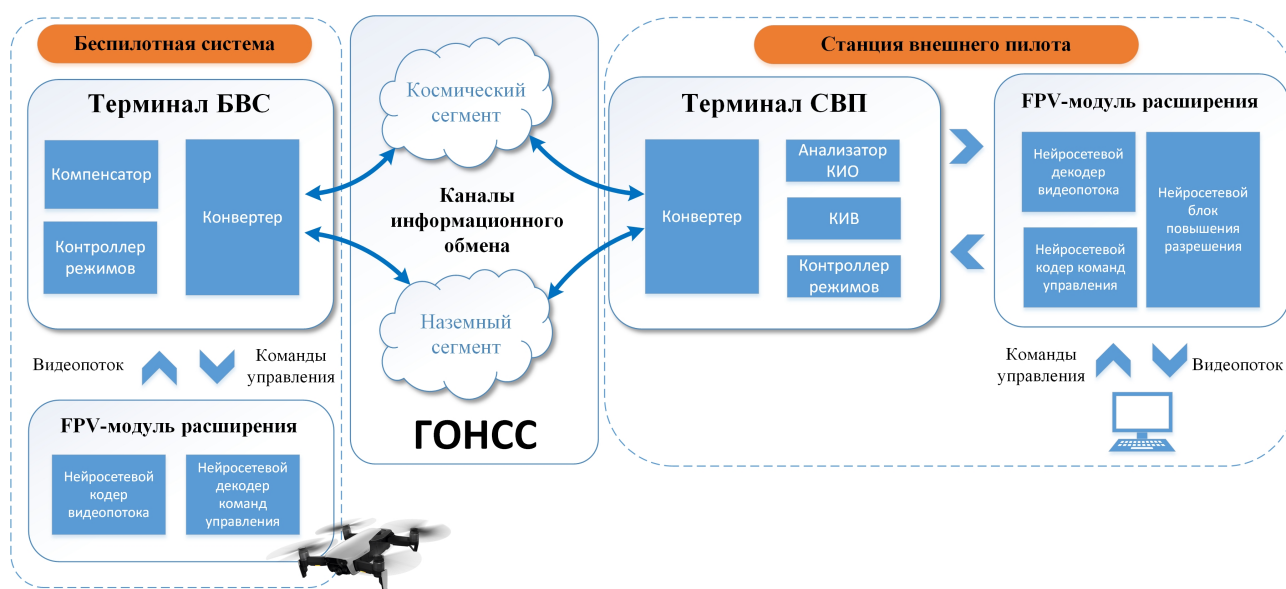


Рис. 1. Структура системы информационного обмена для управления БС в гибридной орбитально-наземной сети связи

Функционально управление потоком данных в каналах информационного обмена на стороне БВС и формирование видеопотока и обработка команд управления разделены на два блока:

- **Терминал БВС.** Назначение терминала БВС заключается в поддержке протокола FPV-STVP (FPV Command Telemetry Video Protocol – протокол передачи команд управления, телеметрии и видео при управлении от первого лица), разрабатываемый в Санкт-Петербургском государственном университете телекоммуникаций им. проф. М. А. Бонч-Бруевича. Терминал БВС принимает закодированный видеопоток от FPV-модуля расширения и отправляет его в сеть на терминал СВП, а также команды управления от СВП и отправляет их на FPV-модуль расширения. Терминал БВС состоит из:
 - конвертера, отвечающего за обработку и формирование сообщений протокола FPV-STVP и управление КИО при передаче через ГОНСС;
 - контроллера режимов, предназначенного для управления режимами (конфигурациями) нейросетевых кодеков видеопотока и команд управления на стороне БВС;
 - компенсатора, обеспечивающего контроль задержки сериализации в КИО [1], а также реализацию функции уменьшения полосы пропускания путём устранения части кадров видеопотока.
- **FPV-модуль расширения.** Его назначение состоит в преобразовании данных в кодированный формат и восстановлении данных из кодированного формата. FPV-модуль расширения получает видеопоток с камеры БС, кодирует и передаёт в терминал БВС, а также

получает кодированные команды управления, восстанавливает их и передаёт в БС в соответствии со спецификацией используемого полетного контроллера. Модуль содержит:

- нейросетевой кодер, который кодирует исходные кадры видеопотока с камеры в сжатую бинарную последовательность в соответствии с конфигурацией [10];
- нейросетевой декодер команд управления, восстанавливающий поступающие с терминала БВС команды управления.

В свою очередь, функционально управление потоком данных в каналах информационного обмена на стороне СВП и обработка видеопотока и формирование команд управления разделены также на два блока:

- **Терминал СВП.** Предназначен для контроля взаимодействия с беспилотной системой по протоколу FPV-CTVP с необходимой интенсивностью видеопотока и синхронизации FPV-управления. Он содержит:

- конвертер, объединяющий функции модема и конвертера терминала БВС: получение и передачу данных через КИО ГОНСС, а также обработку и формирование сообщений протокола FPV-CTVP;
- контроллер режимов, предназначенного для управления режимами (конфигурациями) нейросетевых кодеков видеопотока и команд управления на стороне СВП;
- анализатор КИО, предназначенный для контроля скорости и задержек в канале информационного обмена, а также управление параметрами и компенсационными механизмами КИВ;
- контроллер интенсивности видеопотока, обеспечивающий компенсационные механизмы: увеличение интенсивности видеопотока при его снижении на СВП и уменьшение рассинхронизации FPV-управления.

- **FPV-модуль расширения.** Предназначен для восстановления кадров видеопотока и кодирование (агрегации) команд управления. В отличие от FPV-модуля расширения БВС, на СВП добавляется функциональность повышения качества и разрешения восстановленных кадров видеопотока. Блок состоит из элементов:

- нейросетевой декодер видеопотока, восстанавливающий кадры видеопотока из сжатой бинарной последовательности с применением нейронных сетей;
- нейросетевой кодер команд управления, агрегирующий команды управления от ВП во внутреннее латентное пространство признаков (ЛПП) нейросетевого автокодировщика и на выходе передающий на терминал СВП их сжатое бинарное представление;
- нейросетевой блок повышения разрешения, предназначенный для качественного восстановления разрешения кадров с использованием нейронных сетей, устраняющий артефакты, возникающие в результате кодирования и декодирования кадров нейросетевыми кодерами и декодерами.

Существует несколько особенностей реализации данной системы:

1. В [9, 10] был представлен классический нейросетевой кодек, который включает уменьшение и увеличение разрешения изображения как части своих блоков (блоки интерполяции). В самом исследовании на нейросетевом кодере разрешение кадров уменьшается до 512x512 методом интерполяции суперсэмплингом, а на декодере разрешение кадров восстанавливается до исходного (HD – 1280x720) методом бикубической интерполяции. Тем не менее, в данных блоках могут применяться и другие нейросетевые модели [16]: MagicVideo с латентной диффузией (Latent Diffusion Model, LDM) [11]; SR3 (super resolution 3) [12]; SwinIR (Swin Image Restoration – восстановление изображения с трансформером Swin [13]) [14]; ESRGAN (enhanced super resolution generative adversarial network – генеративно-сопоставительная нейронная сеть для улучшенного сверхразрешения) [15]. В конфигурации, рассматриваемой в настоящей работе, блок интерполяции декодера вынесен отдельно как блок повышения разрешения.

2. Для подавления шума в кадрах видеопотока может применяться специальная версия нейросетевого кодера, реализованная на базе нейросетевой модели латентной диффузии [17, 18]. Однако, данная конфигурация кодера имеет ограниченную эффективность, обеспечивает сжатие кадров до ~ 14 Кб, имеет большое время выполнения с отсутствием адаптивности. В отличие от диффузионного кодера, базовые конфигурации нейросетевого кодера, построенного на базе системы вариационных автокодировщиков [9, 10], имеют лучшие характеристики:

- адаптивное сжатие видеопотока (14 конфигураций) со средней степенью сжатия от $\sim 0,5$ Кб до $\sim 8,5$ Кб;
- меньшее время кодирования и декодирования.

На текущий момент ведутся исследования и разработки по внедрению в нейросетевой кодек новых блоков: подавления артефактов [19], бинаризация/дебинаризация, маскирование/демаскирование [20].

3. Из результатов исследования [2] следует, что предсказание кадров в ЛПП неэффективно с точки зрения работы в реальном времени: кодирование и восстановление кадров занимает намного больше времени, чем в пиксельном пространстве. Соответственно, количество проведения операции декодирования должно быть минимизировано.
4. Как показало исследование [21], значения тензоров ЛПП имеют меньшую автокорреляцию, чем значения кадров видеопотока в пиксельном пространстве. Предсказание тензоров ЛПП менее точно, чем предсказание кадров видеопотока. Соответственно, с точки зрения качества необходимо предсказывать кадры видеопотока именно в пиксельном пространстве, уже после восстановления нейросетевым декодером.
5. Предсказание кадров с помощью специализированной нейронной сети занимает большое количество времени [2] в отличие от механизма интерполяции. Для того, чтобы механизм увеличения интенсивности видеопотока оказывал меньшее воздействие на рассинхронизацию видеопотока, достаточно предсказывать последний кадр с интерполяцией остальных.
6. Кроме использования для получения и передачи данных по сети UDP-сокетов (user datagram protocol – протокол пользовательский датаграмм) [22], имеется поддержка альтернативных вариантов транспортных протоколов FPV-CTVP: TCP (transmission control protocol – протокол управления передачей) [23], QUIC [24, 25].

Система адаптивного контроля интенсивности видеопотока и задержек в КИО входит в модуль «Терминал СВП» (рисунок 1) в составе трёх блоков: контроллер режимов, анализатор КИО и КИВ. Данная система является предметом исследования настоящей работы.

3. Блок контроля интенсивности видеопотока

Блок КИВ – один из блоков системы адаптивного контроля задержки и интенсивности видеопотока (рисунок 2) [7].

Блок КИВ состоит из ряда функциональных блоков.

- Первичный обработчик кадров. Данный блок получает сжатые бинарные представления и метаданные кадров видеопотока и восстанавливает их с помощью нейросетевого декодера. Далее он передаёт восстановленные кадры с их последовательными номерами в блок параметрического учёта кадров. Кроме того, определяет интенсивность видеопотока *FPS*.
- Буфер предсказания - промежуточный буфер кадров для предсказания новых кадров видеопотока.
- Модели предсказания - совокупность всех используемых и поддерживаемых на данном



Рис. 2. Структура блока контроля интенсивности видеопотока

СВП моделей предсказания кадров видеопотока. По-умолчанию используется нейросетевая модель DMVFN. В работе [26] было отмечено высокое быстродействие данной модели, что делает её подходящей для работы в потоковом режиме.

- Блок предсказания кадров. Назначением данного блока является предсказание заданного количества кадров видеопотока из исходных кадров, находящихся в буфере предсказания. Исходные кадры и требование к количеству предсказываемых кадров поступают из блока параметрического учёта кадров. Блок предсказания кадров использует одну из моделей блока «Модели предсказания» в зависимости от настройки терминала СВП.
- Буфер интерполяции - буфер кадров для интерполяции промежуточных кадров видеопотока.
- Модели интерполяции. Совокупность всех используемых и поддерживаемых на данном конкретном СВП моделей интерполяции кадров видеопотока.
- Блок интерполяции кадров. Функциональным назначением данного блока является интерполяция промежуточных кадров видеопотока из исходных кадров, находящихся в буфере интерполяции, с помощью одной из моделей интерполяции (в зависимости от настроек терминала СВП). Исходные кадры и требование к количеству интерполируемых кадров поступают из блока параметрического учёта кадров.
- Блок учета параметров КИВ (рисунок 3). Данный блок предназначен для хранения объема выборки, дисперсии и средних значений параметров КИВ (текущей интенсивности видеопотока, времени интерполяции и предсказания, параметров динамики видеопотока) и предоставления их точечных и интервальных оценок с заданным уровнем значимости. Данные в блоке регулярно обновляются по мере накопления статистики поступления кадров видеопотока.

При заданном уровне значимости α интервальная оценка некоторого параметра PRM вычисляется по формуле [29]

$$PRM \in [PRM_{\min}; PRM_{\max}] = \left[\overline{PRM} - t_{\alpha} \frac{\sqrt{D[PRM]}}{\sqrt{N}}; \overline{PRM} + t_{\alpha} \frac{\sqrt{D[PRM]}}{\sqrt{N}} \right],$$

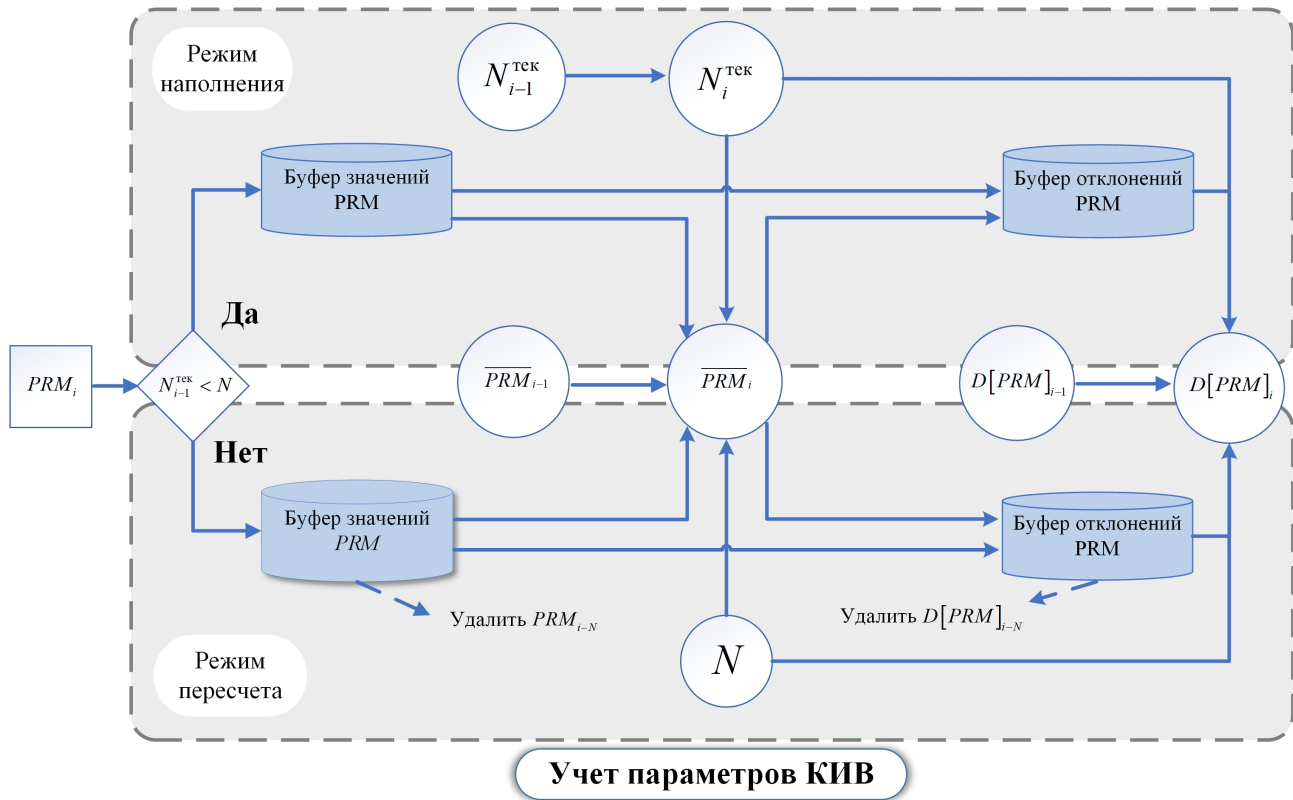


Рис. 3. Принцип работы блока учета параметров КИВ

где $\overline{PRM}$ – среднее значение параметра, $D[PRM]$ – дисперсия параметра, N – объём выборки значений параметров, t_a – коэффициент доверия при заданном уровне значимости a , который рассчитывается по формуле [29]

$$\Phi(t_a) = \frac{1-a}{2},$$

где Φ – функция Лапласа.

Уровень значимости a задаётся настройками терминала СВП при инициализации. По умолчанию уровень значимости $a = 0,05$. Блок учета параметров КИВ хранит до N последних значений каждого из параметров PRM и их квадратов отклонений от средних значений. Размер выборки N задаётся настройками терминала СВП при инициализации. По-умолчанию $N = 30$. Блок «Учет параметров КИВ» имеет два режима работы: режим наполнения и режим пересчёта.

1. Режим наполнения. Пусть в блоке учёта параметров КИВ находится $N_{i-1}^{тек}$ значений параметра PRM , и поступило новое значение параметра PRM_i . Блок учета параметров КИВ находится в режиме наполнения тогда и только тогда, когда $N_{i-1}^{тек} < N$.

После поступления нового значения PRM_i блок учета параметров КИВ пересчитывает количество значений $N_i^{тек}$, среднее значение поступившего параметра $\overline{PRM}$ и дисперсию параметра $D[PRM]$. Новое количество значений $N_i^{тек}$ становится равным $N_i^{тек} = N_{i-1}^{тек} + 1$.

Пусть $\overline{PRM}_{i-1}$ – среднее значение параметра PRM до поступления нового значения PRM_i , а $\overline{PRM}_i$ – обновлённое после поступления. Блок учета параметров КИВ хранит значения параметра PRM

$$\{PRM_{i-N_{i-1}^{тек}}, PRM_{i-N_{i-1}^{тек}+1}, \dots, PRM_{i-1}\}.$$

Тогда обновлённое среднее значение $\overline{PRM}_i$ находится как [30]

$$\overline{PRM}_i = \frac{\overline{PRM}_{i-1} \times N_{i-1}^{\text{тек}} + PRM_i}{N_i^{\text{тек}}}.$$

После обновления среднего значения в блок учета параметров КИВ в режиме наполнения добавляется новое значение параметра PRM_i . После этого он содержит значения:

$$\{PRM_{i-N_i^{\text{тек}}}, PRM_{i-N_i^{\text{тек}}+1}, \dots, PRM_i\}.$$

Если блок учета параметров КИВ в режиме наполнения хранит $N_i^{\text{тек}}$ значений параметра PRM , то перед обновлением дисперсии он хранит $N_{i-1}^{\text{тек}}$ квадратов отклонений параметров от средних значений

$$\left\{ \left(PRM_{i-N_{i-1}^{\text{тек}}} - \overline{PRM}_{i-N_{i-1}^{\text{тек}}} \right)^2, \dots, \left(PRM_{i-1} - \overline{PRM}_{i-1} \right)^2 \right\}$$

и предыдущее значение дисперсии $D[PRM]_{i-1}$.

Тогда обновлённое значение дисперсии $D[PRM]_i$ рассчитывается как

$$D[PRM]_i = \frac{D[PRM]_{i-1} \times N_{i-1}^{\text{тек}} + (PRM_i - \overline{PRM}_i)^2}{N_i^{\text{тек}}}.$$

После обновления значения дисперсии $D[PRM]_i$ в блок учета параметров КИВ в режиме накопления добавляется новое значение квадрата отклонения $(PRM_i - \overline{PRM}_i)^2$. После этого блок учета параметров КИВ содержит значения отклонения:

$$\left\{ \left(PRM_{i-N_i^{\text{тек}}} - \overline{PRM}_{i-N_i^{\text{тек}}} \right)^2, \dots, \left(PRM_i - \overline{PRM}_i \right)^2 \right\}$$

.

2. Режим обновления. Пусть в блоке учета параметров КИВ находится $N_{i-1}^{\text{тек}}$ значений параметра PRM , и поступило новое значение параметра PRM_i . Блок учета параметров КИВ находится в режиме обновления тогда и только тогда, когда $N_{i-1}^{\text{тек}} = N$.

В этом режиме текущее количество значений перестаёт изменяться:

$$N_i^{\text{тек}} = N_{i-1}^{\text{тек}} = N = \text{Const},$$

где Const – константа.

На момент поступления значения нового параметра PRM_i блок учета параметров КИВ в режиме обновления хранит значения PRM

$$\{PRM_{i-N}, PRM_{i-N+1}, \dots, PRM_{i-1}\}$$

и квадраты отклонений PRM

$$\left\{ \left(PRM_{i-N} - \overline{PRM}_{i-N} \right)^2, \dots, \left(PRM_{i-1} - \overline{PRM}_{i-1} \right)^2 \right\}.$$

Обновлённое среднее значение $\overline{PRM}_i$ вычисляется как [30]:

$$\overline{PRM}_i = \frac{\overline{PRM}_{i-1} \times N - PRM_{i-N} + PRM_i}{N} = \overline{PRM}_{i-1} + \frac{PRM_i - PRM_{i-N}}{N}$$

.

После получения обновлённого среднего значения $\overline{PRM}_i$ из блока учета параметров КИВ в режиме обновления удаляется старое значение PRM_{i-N} и добавляется новое значение PRM_i . Таким образом, после этого блок учета параметров КИВ содержит значения параметров PRM :

$$\{PRM_{i-N+1}, PRM_{i-N+2}, \dots, PRM_i\}.$$

Обновлённая дисперсия $D[PRM]_i$ вычисляется как

$$D[PRM]_i = D[PRM]_{i-1} + \frac{(PRM_i - \overline{PRM}_i)^2 - (PRM_{i-N} - \overline{PRM}_{i-N})^2}{N}$$

После получения обновлённой дисперсии $D[PRM]_i$ из блока учета параметров КИВ в режиме пересчёта удаляется устаревшее значение квадрата отклонения $(PRM_{i-N} - \overline{PRM}_{i-N})^2$ и добавляется новое значение $(PRM_i - \overline{PRM}_i)^2$. Таким образом, после этого блок учета параметров КИВ содержит значения квадратов отклонений:

$$\{(PRM_{i-N+1} - \overline{PRM}_{i-N+1})^2, \dots, (PRM_i - \overline{PRM}_i)^2\}.$$

- Текущая интенсивность видеопотока - функциональная часть блока «Учет параметров КИВ», необходимая для получения точечных и интервальных оценок текущей интенсивности видеопотока в соответствии с заданным уровнем значимости (в зависимости от настроек терминала СВП). Хранит, обновляет и предоставляет среднее значение, дисперсию и объём выборки интенсивности видеопотока (FPS). FPS измеряется в кадрах в секунду (кадр/с).
- Динамика видеопотока - функциональная часть блока «Учет параметров КИВ», необходимая для получения точечных и интервальных оценок текущей динамики видеопотока в соответствии с заданным уровнем значимости (в зависимости от настроек терминала СВП). Хранит, обновляет и предоставляет среднее значение, дисперсию и объём выборки динамики видеопотока DYN . DYN представляет собой значение метрики $PSNR$ между кадрами видеопотока, отражающими полётную обстановку, отличающуюся на 40 мс (с учетом $FPS = 25$). DYN измеряется между кадрами на временном промежутке:

$$ddyn \approx \left\lceil \frac{40}{\left(\frac{1000}{FPS}\right)} \right\rceil = \left\lceil \frac{40 \times FPS}{1000} \right\rceil :$$

$$DYN \approx PSNR(x_{i-ddyn}, x_i) \text{ (дБ)},$$

где x_i – текущий кадр, x_{i-ddyn} – прошлый кадр под номером $i - ddyn$, i – порядковый номер, $PSNR(x_i, x_{i-ddyn})$ – пиковое отношение сигнала к шуму (peak sound-to-noise ratio) между кадрами x_i и x_{i-ddyn} , рассчитывающееся как [27]

$$PSNR(x_i, x_{i-ddyn}) = 20 \times \log_{10} \left(\frac{255}{\sqrt{MSE(x_i, x_{i-ddyn})}} \right),$$

где $MSE(x_i, x_{i-ddyn})$ – средний квадрат ошибки (mean squared error) между кадрами x_i и x_{i-ddyn} , вычисляемый в соответствии с выражением [27, 28]:

$$MSE(x, x_{i-ddyn}) = \frac{1}{height \times width} \sum_{h=0}^{height} \sum_{w=0}^{width} (x_i(h, w) - x_{i-ddyn}(h, w))^2,$$

где $height$ – высота кадров в пикселях x_i и x_{i-ddyn} , а $width$ – ширина кадров в пикселях, $x_i(h, w)$ – значение пикселя $(h \times w)$ текущего кадра x_i , а $x_{i-ddyn}(h, w)$ – значение пикселя $(h \times w)$ кадра x_{i-ddyn} .

- Время интерполяции - функциональная часть блока «Учет параметров КИВ», необходимая для получения точечных и интервальных оценок времени интерполяции в соответствии с заданным уровнем значимости (в зависимости от настроек терминала СВП). Хранит и обновляет: среднее значение, дисперсию и объём выборки времени интерполяции. Измеряется в миллисекундах (мс).
- Время предсказания - функциональная часть блока «Учет параметров КИВ», необходимая для получения точечных и интервальных оценок времени предсказания (одного кадра) в соответствии с заданным уровнем значимости (в зависимости от настроек терминала СВП). Хранит и обновляет: среднее значение, дисперсию и объём выборки времени предсказания одного кадра. Измеряется в миллисекундах (мс).
- Параметрический учёт кадров. Данный блок формирует буфер кадров видеопотока для адаптивных режимов работы: *режим увеличения интенсивности видеопотока* и *режим уменьшения задержки видеопотока*. Для реализации данных режимов использует параметры из блока учета параметров КИВ и восстановленные кадры видеопотока из блока первичной обработки кадров. Для поддержки своих режимов работы использует результаты работы блока предсказания кадров видеопотока и блока интерполяции кадров. Функциональность блока параметрического учёта кадров делится на *базовую* и *адаптивную*. Подробнее об адаптивной функциональности см. разделы «Режим увеличения интенсивности видеопотока» и «Режим уменьшения задержки видеопотока».

Базовую функциональность (рисунок 4) можно описать следующим алгоритмом:

- От блока «Первичный обработчик кадров» поступают восстановленный кадр видеопотока x_k , его порядковый номер k и параметр текущей интенсивности видеопотока FPS .
- Кадр видеопотока x_k поступает в буфер учета кадров B_X , из которого удаляется последний неактуальный кадр x_{i-N+1} .
- Далее рассчитывается разница между номерами текущего и последнего кадров $\Delta = k - i$.
После этого значение i замещается значением k . Номер кадра k поступает в буфер номеров B_N , из которого удаляется номер $i - N + 1$.
- Если $\Delta = 1$ или $\Delta > N_{пред}^{инт}$, то в модуль «Отображение видеопотока» поступает только кадр x_k .
- Если $\Delta \leq N_{пред}^{инт}$, где $N_{пред}^{инт}$ – параметр текущей модели интерполяции (максимальное число кадров, интерполируемых с приемлемым качеством), то из буфера учета кадров B_X извлекается кадр x_i (до замещения i на k). Кадры x_i и x_k , а также значение Δ (количество интерполируемых кадров) отправляются в блок интерполяции кадров. На отображение оператору отправляются как полученный на выходе буфер интерполированных кадров B_I , так и поступивший кадр видеопотока x_k .
- Если кадр x_{k-ddyn} присутствует в буфере кадров B_X , то рассчитывается значение $PSNR$ между кадрами x_k и x_{k-ddyn} . На основе значения метрики $PSNR$ и интенсивности видеопотока FPS рассчитывается динамика видеопотока DYN . Обновленное значение параметра DYN отправляется в блок учета параметров КИВ.
- Значение текущей интенсивности видеопотока FPS отправляется в блок «Учет параметров КИВ». После этого вычисляется минимальное значение $FPS_{мин}$ интервальной оценки FPS из блока «Учет параметров КИВ». Минимальная текущая интенсивность кадров $FPS_{мин}$ отправляется в блок «Отображение видеопотока».
- От блока «Интерполяция кадров видеопотока» поступают: время интерполяции T_I ,

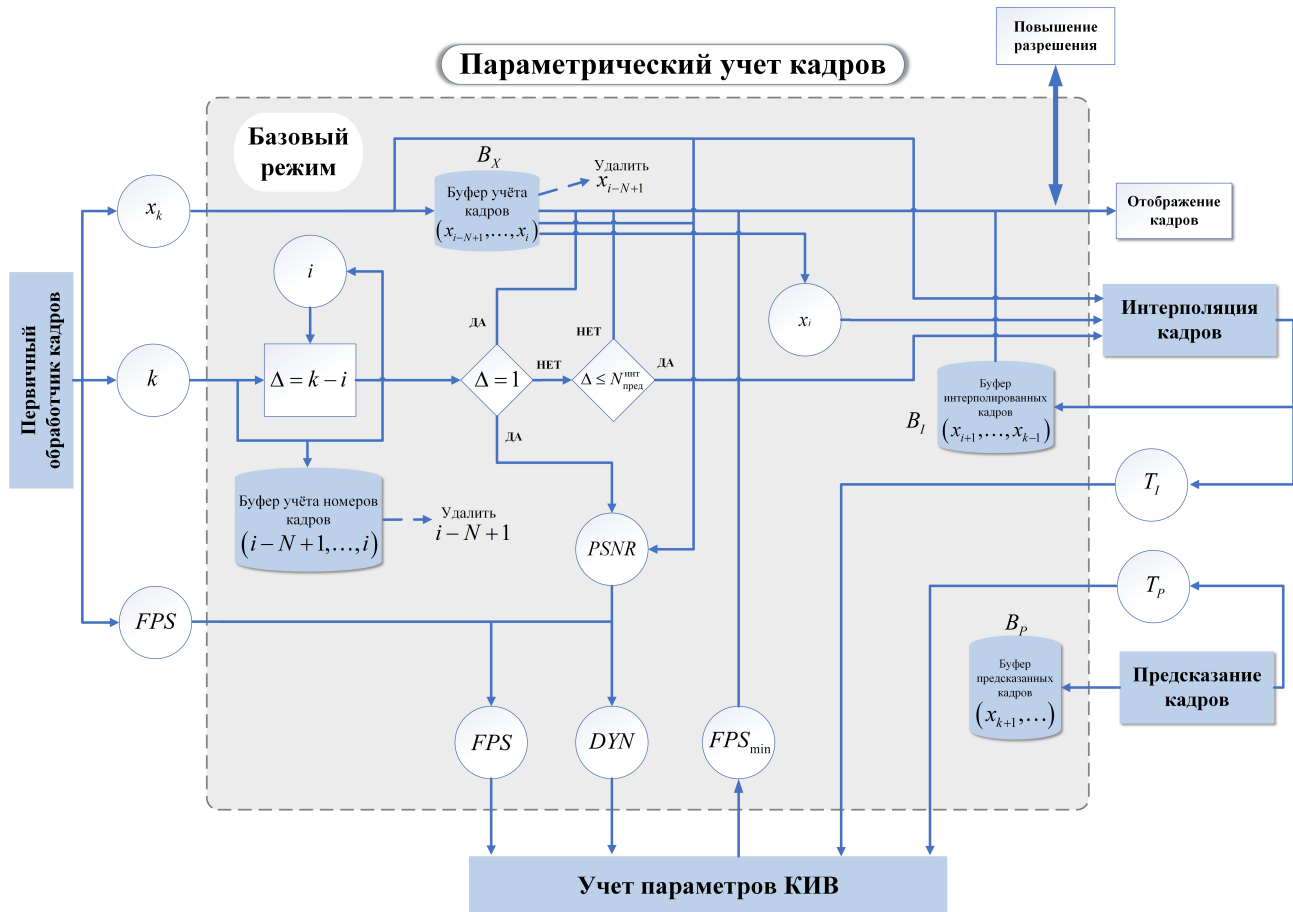


Рис. 4. Базовая функциональность блока параметрического учёта кадров

буфер интерполированных кадров B_I . Время интерполяции T_I отправляется в блок «Учет параметров КИВ». Кадры из буфера интерполированных кадров отправляются оператору в блок «Отображение видеопотока».

- От блока «Предсказание кадров видеопотока» поступает время предсказания T_P и буфер предсказанных кадров B_P . Время предсказания T_P отправляется в блок «Учет параметров КИВ». Блок «Предсказание кадров видеопотока» не применяется в базовом режиме, но функциональность учёта поступающего времени предсказания относится к базовой.
- Перед отправкой в блок «Отображение видеопотока» разрешение всех кадров видеопотока восстанавливается в соответствии с настройкой терминала СВП в блоке «Повышение разрешения». По умолчанию восстанавливается разрешение HD (1280x720).

4. Анализатор канала информационного обмена

Блок контроля интенсивности видеопотока (КИВ) (рисунок 2) связан с блоком анализатора канала информационного обмена (КИО) (рисунок 5). Он содержит ряд функциональных подблоков:

- Учёт параметров КИО. Блок, который получает значения, хранит, предоставляет точечные и интервальные оценки параметра КИО – прикладной задержки видеопотока T_3^B . Принцип работы блока учёта параметров КИО для прикладной задержки идентичен принципу работы блока учёта параметров КИВ (рисунок 3). Также хранит точечные значения



Рис. 5. Структура блока анализатора канала информационного обмена

параметров: текущий режим, количество кадров для предсказания, количество кадров для интерполяции.

- Прикладная задержка видеопотока T_3^B – задержка передачи кадров видеопотока в сжатом бинарном представлении, измеряется в мс.
- Текущий режим КИВ (базовый, увеличения интенсивности видеопотока, уменьшения задержки).
- Количество кадров для предсказания $N_{\text{пред}}$. Параметр режима КИВ по уменьшению задержки видеопотока. Хранит количество кадров, на которые необходимо выполнять предсказание.
- Количество кадров для интерполяции $N_{\text{инт}}$. Параметр режима увеличения интенсивности видеопотока. Хранит количество промежуточных кадров, на которые необходимо интерполировать.
- Блок измерения прикладной задержки. Каждые $PROB$ секунд блок измерения прикладной задержки отправляет через внешний блок «Конвертер» сообщение об измерении прикладной задержки в рамках протокола FPV–CTVP. Длина сообщения равна средней длине сжатых бинарных последовательностей кадров для используемой конфигурации нейросетевого кодека. Полученное значение прикладной задержки T_{zi}^B отправляется в блок учёта параметров КИО. Значение $PROB$ зависит от настройки терминала СВП. По умолчанию $PROB = 5$ сек.
- Монитор параметров КИО. Задача данного блока заключается в обнаружении аномальных изменений параметров интенсивности видеопотока FPS (рисунок 6) [7] и прикладной задержки видеопотока T_3^B . Период мониторинга MON определяется настройками терминала СВП и по умолчанию $MON = 60$ сек.

- Модели монитора параметров КИО. Множество всех поддерживаемых на конкретном СВП и используемых в мониторе параметров КИО моделей. Модели монитора параметров КИО используются для анализа временных рядов изменения прикладной задержки и интенсивности видеопотока, а также обнаружения аномалий в рядах. В качестве моделей монитора параметров КИО могут быть использованы [7, 31] модели ChangePointDetectionClassifier, ChangePointDetectionClassifierCV, OnlineNNClassifier, SlidingWindows, ChangePoint-DetectionRuLSIF, OnlineNNRuLSIF), Skoltech Anomaly Benchmark (SKAB) [32] и статистические карты Шухарта [33].

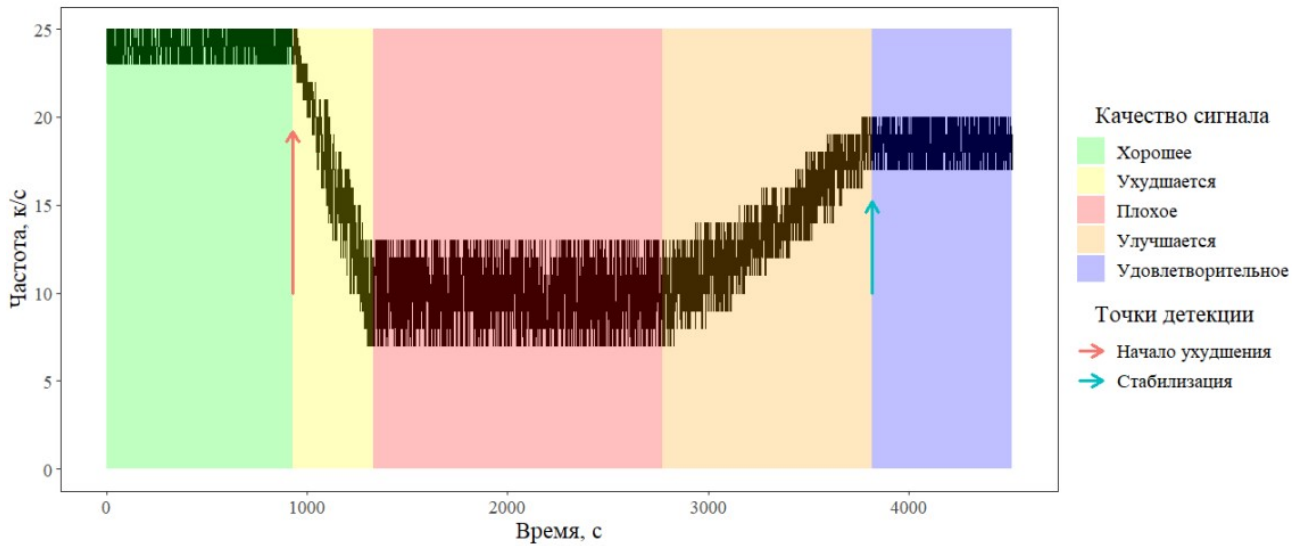


Рис. 6. Аномалии интенсивности видеопотока

При возникновении аномального уменьшения значения FPS , то включается режим увеличения интенсивности видеопотока, который отключается при нормализации FPS до установленных значений. При возникновении аномального увеличения задержки видеопотока, например, при смене сегментов передачи видеопотока, то включается режим уменьшения задержки. При аномальном уменьшении прикладной задержки видеопотока, режим уменьшения задержки отключается. При этом конфигурация нейросетевого кодека во внешнем блоке «Контроллер режимов» сбрасывается до конфигурации с максимально возможным качеством восстанавливаемых кадров видеопотока.

Режимы являются взаимоисключающими. В одно время может быть включен либо режим увеличения интенсивности, либо режим уменьшения задержки. При этом в случае с базовым режимом работы блока параметрического учета кадров, адаптивные режимы «анализатора КИО» отключены.

- Адаптивный контроллер интенсивности. Данный блок КИО задействуется в адаптивном режиме увеличения интенсивности видеопотока. Пусть интенсивность предыдущего процесса интенсивности равна FPS_{i-1} , а текущего — FPS_i . Интерполяция одного кадра увеличивает интенсивность видеопотока в два раза, двух — в три раза и так далее. Число интерполируемых кадров $N_{\text{инт}}$ подбирается таким образом, чтобы интенсивность предыдущего процесса была гарантированно достигнута:

$$(N_{\text{инт}} + 1) \times FPS_i = FPS_{i-1},$$

$$N_{\text{инт}} = \frac{FPS_{i-1}}{FPS_i} - 1.$$

Число интерполируемых кадров обновляется в зависимости от предельного числа приемлемо интерполируемых кадров используемой модели интерполяции:

$$N_{\text{инт}} = \min(N_{\text{инт}}; N_{\text{пред}}^{\text{инт}}).$$

Число интерполируемых кадров $N_{\text{инт}} > 0$ (так как по определению режима увеличения интенсивности видеопотока $FPS_{i-1} > FPS_i$) не обязательно должно являться целым числом (подробнее см. раздел «Режим увеличения интенсивности видеопотока»).

Пересчёт числа интерполируемых кадров $N_{\text{инт}}$ проводится каждые MON секунд.

- Адаптивный контроллер задержек КИО. Данный блок КИО задействуется в адаптивном режиме уменьшения задержки распространения. Его назначением является уменьшение прикладной задержки видеопотока и, таким образом, времени рассинхронизации FPV-управления. В начале контроллер задержек пытается достичь уменьшения прикладной задержки, последовательно выбирая через блок «Контроллер режимов» конфигурацию нейросетевого кодека с большим коэффициентом сжатия кадров и, следовательно, худшим качеством. При этом, если достигнута конфигурация с минимальным размером сжатого кадра на выходе и не происходит уменьшения прикладной задержки, то выбирается та конфигурация нейросетевого кодека, после которой не было статистически значимого уменьшения задержки.

Пусть $T_{zi(k)}^B$ – прикладная задержка текущего процесса на k -конфигурации, и порядковый номер конфигураций начинается с 1. Тогда выбирается конфигурация под таким номером k , что

$$\forall q \in [1; k-1] : T_{zi(k)}^B \leq T_{zi(k-q)}^B.$$

Задержка предыдущего процесса T_{zi-1}^B , а текущего с конфигурацией $k - T_{zi}^B$. Если условие $T_{zi}^B > T_{zi-1}^B$ не выполняется, то режим уменьшения задержки отключается без изменения конфигурации. Если выполняется, то рассчитываются параметры предсказания кадров видеопотока:

$$\Delta T_3^B = T_{zi}^B - T_{zi-1}^B \text{ (мс)},$$

где ΔT_3^B – время предсказания полетной обстановки вперед.

Исследование [1] показало, что при высокоинтенсивном видеопотоке возможно предсказание на 40 мс вперёд, а при низкоинтенсивном – на 80-120 мс и более. По результатам [1] эмпирически подбирается значение предельной дальности предсказания $\Delta T_{3 \text{ макс}}^B$ (таблица 1).

Т а б л и ц а 1. Таблица предельной дальности предсказания

DYN , дБ	$\Delta T_{3 \text{ макс}}^B$, мс
(36; ∞)	120
(32; 36]	80
(30; 32]	40
(0; 30]	0

После этого значение ΔT_3^B обновляется:

$$\Delta T_3^B = \min(\Delta T_3^B; \Delta T_{3 \text{ макс}}^B).$$

Число кадров опережения при предсказании рассчитывается как

$$N_{\text{пред}} = \left\lceil \frac{\Delta T_3^{\text{в}}}{FPS_{\text{мин}}} \right\rceil.$$

В отличие от числа интерполируемых кадров $N_{\text{инт}}$, число кадров опережения при предсказании обязательно является натуральным числом, т.е. $N_{\text{пред}} \in \mathbb{N}$.

Пересчёт конфигурации k и числа кадров опережения при предсказании $N_{\text{пред}}$ проводится каждые MON секунд.

5. Режим увеличения интенсивности видеопотока

При включении адаптивного режима увеличения интенсивности видеопотока блока КИВ «Параметрический учёт кадров» добавляется адаптивная функциональность увеличения интенсивности видеопотока (рис. 7). В данном режиме в блок «Параметрический учёт кадров» дополнительно поступает параметр количества интерполируемых кадров $N_{\text{инт}}$ и добавляется второй буфер интерполяции B_{I2} , где основной буфер интерполяции обозначается как первый B_{I1} .

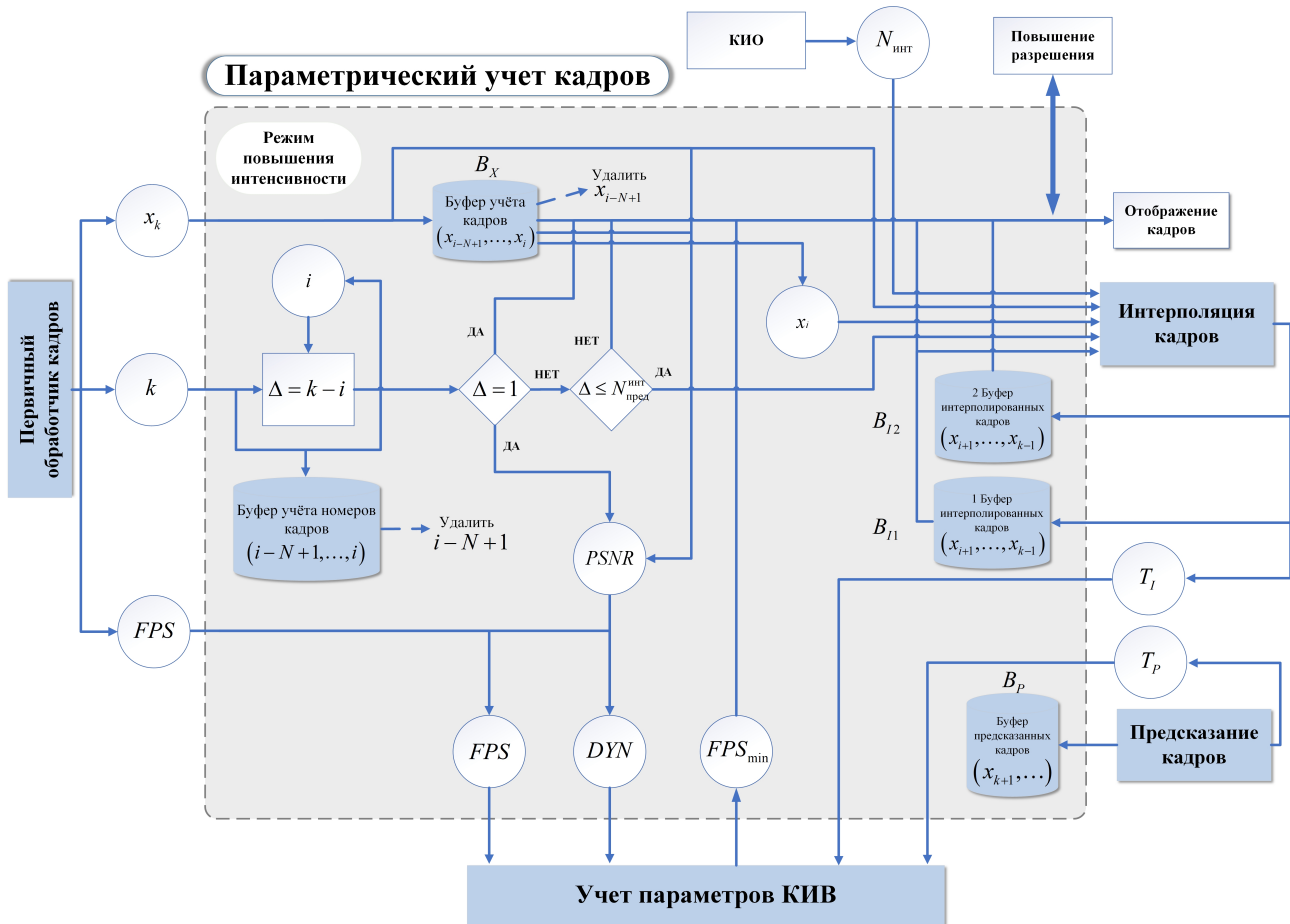


Рис. 7. Функциональность блока параметрического учёта кадров в режиме увеличения интенсивности

Адаптивную функциональность (рис. 7) в части увеличения интенсивности видеопотока можно описать следующим алгоритмом:

- При выполнении одного из условий: $\Delta = 1$ или $\Delta > N_{\text{пред}}^{\text{инт}}$, - в адаптивном режиме кадры x_i и x_k поступают в блок «интерполяции кадров» один раз. Между ними интерполиру-

ется $N_{\text{инт}}$ (при $N_{\text{инт}} \in \mathbb{N}$) кадров, которые поступают во второй буфер интерполяции B_{I2} при этом первый буфер интерполяции B_{I1} остаётся пустым.

- Если $\Delta \neq 1$ и $\Delta \leq N_{\text{пред}}^{\text{инт}}$, в адаптивном режиме увеличения интенсивности видеопотока исходные кадры x_i , x_k и интерполированные кадры из первого буфера интерполяции B_{I1} повторно поступают в блок «интерполяции кадров». Между парами соседних по номеру кадров дополнительно интерполируется по $N_{\text{инт}}$ (при $N_{\text{инт}} \in \mathbb{N}$) кадров (рис. 8) [2]. Полученные в результате кадры поступают в B_{I2} .

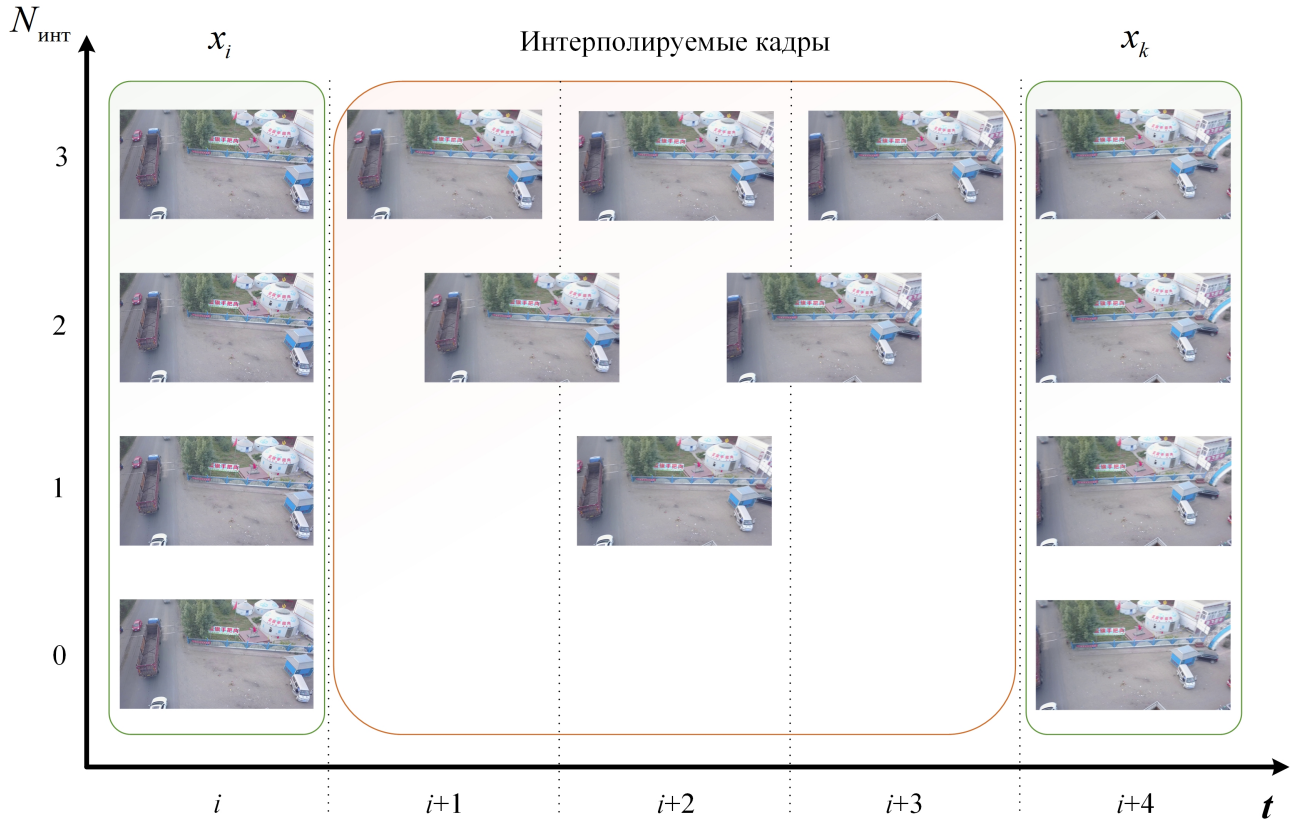


Рис. 8. Интерполяция промежуточных кадров

- Если $N_{\text{инт}} \notin \mathbb{N}$, то количество интерполируемых кадров варьируется от $\lfloor N_{\text{инт}} \rfloor$ до $\lceil N_{\text{инт}} \rceil$. В блоке «Интерполяции кадров» создаётся счётчик количества интерполируемых кадров $INT_COUNTER = 0$. Перед каждой новой интерполяцией счётчик увеличивается на $N_{\text{инт}}$. Количество кадров, которые интерполируются в этот момент $N_{\text{инт}(i)}$, определяется как $N_{\text{инт}(i)} = \lfloor INT_COUNTER \rfloor$. После завершения интерполяции кадров счётчик уменьшается на значение $N_{\text{инт}(i)}$.
- Исходный кадр x_k , кадры из буферов B_{I1} и B_{I2} упорядочиваются по их номерам. Их разрешение восстанавливается блоком «Повышение разрешения». Далее они поступают на блок «Отображение кадров».

Дополнительная задержка, вносимая работой адаптивного режима повышения интенсивности, приблизительно равна

$$\Delta T_{\text{инт}} \approx N_{\text{инт}} \times T_I.$$

6. Режим уменьшения задержки

При включении адаптивного режима уменьшения задержки видеопотока у блока КИВ «Параметрический учёт кадров» добавляется адаптивная функциональность предсказания будущих кадров видеопотока (рис. 9).

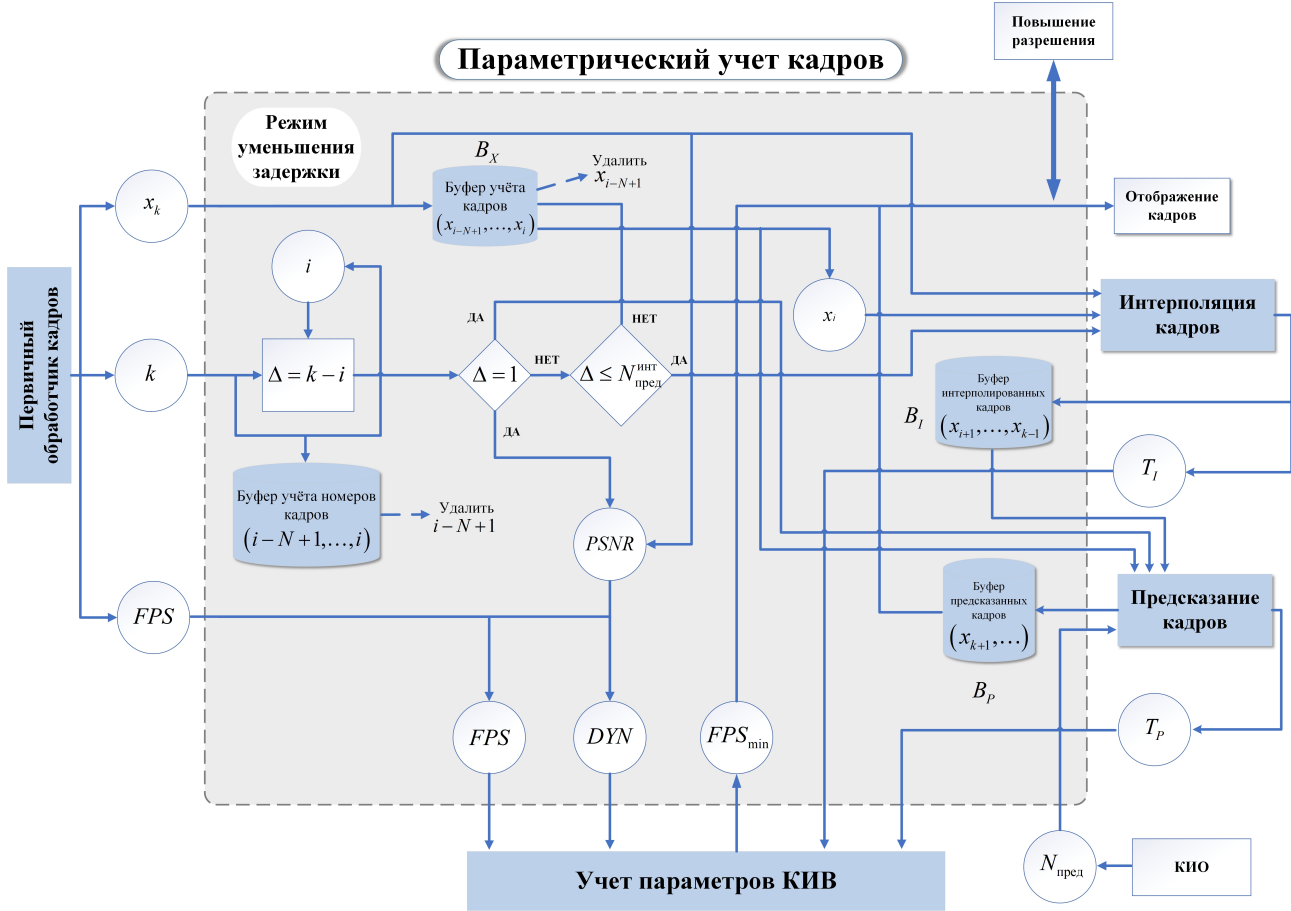


Рис. 9. Функциональность параметрического учёта кадров в режиме уменьшения задержки

Адаптивную функциональность в части предсказания будущих кадров видеопотока можно описать следующим алгоритмом:

- В блок «Параметрический учёт кадров» дополнительно поступает число кадров опережения при предсказании $N_{\text{пред}}$. Данный параметр поступает в функциональный блок «Предсказание кадров».
- При выполнении одного из условий: $\Delta = 1$ или $\Delta > N_{\text{пред}}^{\text{инт}}$, - в адаптивном режиме уменьшения задержки из буфера кадров видеопотока B_X извлекается кадр $x_{k-N_{\text{пред}}}$. Кадры x_k и $x_{k-N_{\text{пред}}}$ поступает в блок «Предсказание кадров». В результате в буфер предсказания B_P поступает предсказанный кадр $\hat{x}_{k+N_{\text{пред}}}$.
- Пусть интерполированным кадрам соответствуют моменты времени $i+1, i+2, \dots, i+\Delta$. Если $\Delta \neq 1$ и $\Delta \leq N_{\text{пред}}^{\text{инт}}$, то в режиме уменьшения задержки из буфера кадров B_X извлекаются кадры

$$x_{i+1-N_{\text{пред}}}, x_{i+2-N_{\text{пред}}}, \dots, x_{i+\Delta-N_{\text{пред}}}.$$

Из буфера кадров B_X и буфера интерполяции B_I в блок предсказания поступают пары кадров $((x_k, x_{k-N_{\text{пред}}}), (x_{i+1}, x_{i+1-N_{\text{пред}}}), (x_{i+2}, x_{i+2-N_{\text{пред}}}), \dots, (x_{i+\Delta}, x_{i+\Delta-N_{\text{пред}}}))$.

Далее из каждой пары блок «Предсказание кадров» предсказывает по одному кадру:

$$x_{i+1+N_{\text{пред}}}, x_{i+2+N_{\text{пред}}}, \dots, x_{i+\Delta-N_{\text{пред}}}, x_{k+N_{\text{пред}}}.$$

Предсказанные кадры $\hat{x}_i$ поступают в буфер предсказанных кадров B_P .

- Разрешение кадров из буфера предсказания B_P увеличивается в блоке повышения разрешения с 512x512 до разрешения видеопотока терминала СВП. Затем данные кадры поступают в блок отображения (рис. 10) [1], где B_D – буфер отображения кадров видеопотока и $\hat{t}_i$ вычисляется в соответствии с выражением:

$$\hat{t}_i = t_i + T_B^3 + T_P.$$

Адаптивный режим уменьшения задержки сокращает рассинхронизацию видеопотока приблизительно на:

$$\Delta T_p \approx N_{\text{пред}} \times FPS_{\text{мин}} - T_P.$$

7. Заключение

В настоящем исследовании рассмотрены факторы, влияющие на уменьшение вероятности успешного достижения цели функционирования БВС при FPV-управлении: низкая интенсивность видеопотока и высокая рассинхронизация FPV-управления. Приведены исследования, в которых предложены методы нивелирования данных проблем по отдельности путём интерполяции промежуточных кадров и предсказания будущих кадров. Подробно рассмотрена функциональность системы информационного обмена для управления БВС, предложенная в ранних работах авторов.

Научная новизна работы заключается в том, что в настоящей работе предлагается структуры системы, обеспечивающей динамическое увеличение интенсивности и задержки видеопотока при изменении задержек и потерь кадров видеопотока в канале связи с использованием нейронных сетей.

Система адаптивного контроля интенсивности и уменьшения задержки видеопотока включает блоки контроллера режимов, контроля интенсивности видеопотока и анализатор канала информационного обмена.

Подробно рассмотрен блок КИВ, назначением которого является контроль за отображением кадров видеопотока ВП и применение адаптивных механизмов для регулирования величины рассинхронизации и интенсивности видеопотока на СВП. КИВ функционирует в трех режимах работы: базовом и двумя адаптивными (увеличения интенсивности видеопотока и уменьшения задержки). В режиме повышения интенсивности видеопотока КИВ интерполирует промежуточные кадры видеопотока, увеличивая интенсивность кадров видеопотока на оборудовании отображения СВП. В режиме уменьшения задержки КИВ предсказывает будущие кадры видеопотока и тем самым уменьшает задержку в КИО (прикладную задержку) для передачи видеопотока и, как следствие, время рассинхронизации FPV-управления. Разработана структура и функциональное содержание данного блока, методика хранения и получения точечных и интервальных оценок значений параметров КИВ (интенсивность видеопотока, динамика видеопотока, время интерполяции и предсказания). Отдельно рассмотрена методика функционирования КИВ в адаптивных режимах.

Также рассмотрен блок анализатора каналов информационного обмена, который контролирует задержку в КИО для передачи видеопотока, анализирует интенсивность и прикладную задержку видеопотока на предмет аномального изменения, управляет работой адаптивных режимов КИВ, рассчитывает параметры адаптивных режимов. Разработана структура и функ-

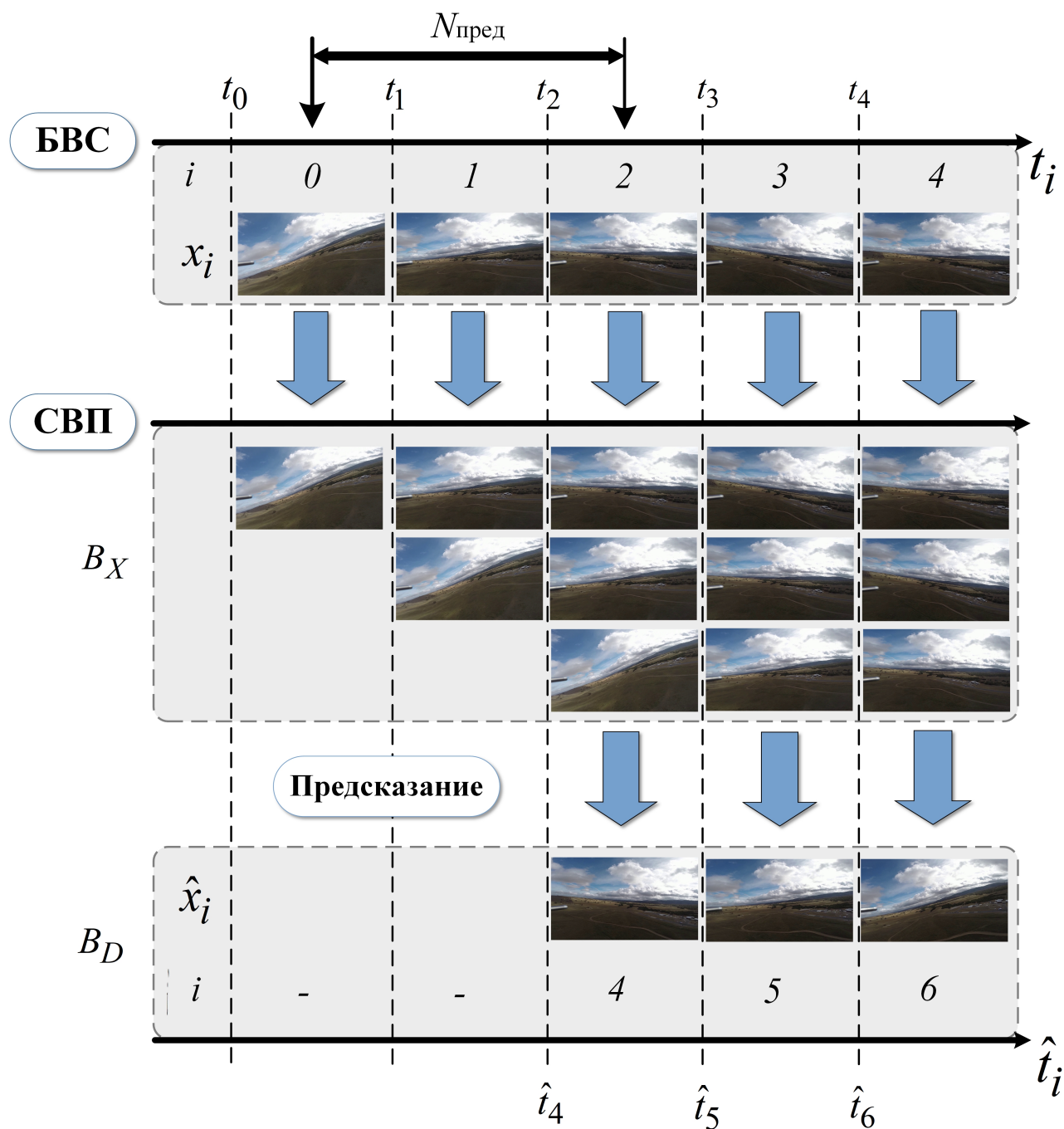


Рис. 10. Предсказание будущих кадров

циональное содержание блока КИВ. Представлена методика расчёта параметров адаптивных режимов.

Система адаптивного КИВ и уменьшения задержки видеопотока может быть использована в составе нейросетевых кодеков для обеспечения стабильного уровня интенсивности видеопотока в наземном сегменте ГОНСС и контроля задержек видеопотока в канале информационного обмена космического сегмента ГОНСС при FPV-управлении БВС.

Литература

1. Березкин А. А., Ченский А. А., Киричек Р. В. Уменьшение задержки передачи FPV-видеопотока через спутниковые каналы связи методом предсказания будущих кадров // *Электросвязь*. 2025. № 4.
2. Березкин А. А., Ченский А. А., Киричек Р. В. Исследование границ интенсивности видеопотока при FPV-управлении БПЛА в режиме предсказания кадров. Часть I: модели и методы // *Вестник СибГУТИ*. 2024. Т. 18. №. 3. С. 115-139.
3. Hu X., Huang Z., Huang A., Xu J., Zhou S. A dynamic multi-scale voxel flow network for video prediction // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023. С. 6121–6131.
4. Choquette J., Gandhi W. Nvidia A100 GPU: Performance & innovation for GPU computing // 2020 IEEE Hot Chips 32 Symposium (HCS). IEEE Computer Society, 2020. С. 1-43.
5. Choquette J., Gandhi W., Giroux O., Stam N., Krashinsky R. Nvidia A100 tensor core GPU: Performance and innovation // *IEEE Micro*. 2021. Т. 41. №. 2. С. 29-35.
6. Giambene G., Kota S., Pillai P. Satellite-5G integration: A network perspective // *IEEE Network*. 2018. Т. 32. №. 5. С. 25-31.
7. Березкин А. А., Паришин А. А., Лазарев А. А. Адаптивный контроль интенсивности видеопотока при передаче FPV-трафика беспилотных систем // *Материалы 79-й научно-технической конференции СПб НТО РЭС им. А. С. Попова, посвященной Дню радио, Санкт-Петербург, 22–26 апреля 2024*. С. 162–165.
8. Стратегия развития отрасли связи Российской Федерации на период до 2035 года. // Правительство России. URL: <http://government.ru/news/50304/> (дата обращения: 02.01.2025).
9. Березкин А. А., Ченский А. А., Киричек Р. В., Захаров А. А. Исследование конфигураций нейросетевых кодеков для адаптивной системы сжатия кадров FPV-видеопотока при управлении беспилотными системами. Часть I. Методика // *Электросвязь*. Санкт-Петербург, 2024. № 9. С. 42-51.
10. Березкин А. А., Ченский А. А., Киричек Р. В., Захаров А. А. Исследование конфигураций нейросетевых кодеков для адаптивной системы сжатия кадров FPV-видеопотока при управлении беспилотными системами. Часть II. Эксперимент // *Электросвязь*. Санкт-Петербург, 2024. № 10. С. 59-69.
11. Zhou D., Wang W., Yan H., Lv W., Zhu Y., Feng J. MagicVideo: Efficient video generation with latent diffusion models. // *arXiv preprint arXiv:2211.11018*. 2022.
12. Saharia C., Ho J., Chan W., Salimans T., Fleet D. J., Norouzi M. Image super-resolution via iterative refinement // *IEEE transactions on pattern analysis and machine intelligence*. 2022. v. 45. № 4. p. 4713–4726.
13. Liu Z., Lin Y., Cao Y., Hu H., Wei Y., Zhang Z., Lin S., Guo B. SWIN transformer: Hierarchical vision transformer using shifted windows // *Proceedings of the IEEE/CVF international conference on computer vision*. 2021. С. 10012-10022.
14. Liang J., Cao J., Sun G., Zhang K., Van Gool L., Timofte R. SwinIR: Image restoration using swin transformer // *Proceedings of the IEEE/CVF international conference on computer vision*. 2021. С. 1833–1844.
15. Wang X., Yu K., Wu S., Gu J., Liu Y., Dong C., Qiao Y., Loy C. C. ESRGAN: Enhanced super-resolution generative adversarial networks // *Proceedings of the European conference on computer vision (ECCV) workshops*. 2018.
16. Козлова А. И., Облаков Н. А., Березкин А. А. Сравнительный анализ методов сверхразрешения в системе сжатия видеопотока при FPV управлении беспилотными системами // *Материалы 79-й научно-технической конференции СПб НТО РЭС им. А. С. Попова, посвященной*

ной Дню радио, Санкт-Петербург, 22–26 апреля 2024. С. 166–169.

17. Березкин А. А., Вивчарь Р. М., Слепнев А. В., Киричек Р. В., Захаров А. А. Метод сжатия видеопотока при управлении беспилотными системами в гибридных орбитально-наземных сетях связи // *Электросвязь*. 2023. №. 10. С. 48–56.
18. Березкин А. А., Вивчарь Р. М., Киричек Р. В., Захаров А. А. Метод декомпрессии FPV-видеопотока от беспилотных систем на основе латентной диффузионной нейросетевой модели // *Электросвязь*. 2024. №. 1. С. 42–53.
19. Березкин А. А., Ченский А. А., Киричек Р. В., Захаров А. А. Маскирование кодированного FPV видеопотока при управлении БПЛА. Часть I. Модели и методика // *Электросвязь*. (Принято к публикации).
20. Березкин А. А., Ченский А. А., Киричек Р. В. Нивелирование артефактов кадров видеопотока при FPV-управлении беспилотными системами // *Инфкоммуникационные технологии*. 2024. Т. 22. № 1 (85). С. 7–17.
21. Березкин А. А., Облаков Н. А., Ченский А. А., Киричек Р. В. Корреляционный анализ латентного видеопотока при управлении БАС от первого лица // *Электросвязь*. Санкт-Петербург, 2024. № 11. С. 25–31.
22. Larzon L. A., Degermark M., Pink S. UDP lite for real time multimedia applications. // Hewlett-Packard Laboratories, 1999.
23. Allman M., Falk A. On the effective evaluation of TCP // *ACM SIGCOMM Computer Communication Review*. 1999. v. 29. №. 5. p. 59–70.
24. Langley A., Riddoch A., Wilk A., Vicente A., Krasic C., Zhang D., Shi Z. The QUIC transport protocol: Design and internet-scale deployment // *Proceedings of the conference of the ACM special interest group on data communication*. 2017. p. 183–196.
25. Cui Y., Li T., Liu C., Wang X., Kühlewind M. Innovating transport with QUIC: Design approaches and research challenges // *IEEE Internet Computing*. 2017. v. 21. №. 2. p. 72–76.
26. Березкин А. А., Матвеев Д. В., Савелов Д. Ю. Оптимизация модели предсказания видео с помощью Tensorrt для уменьшения задержек при управлении БПЛА // *Подготовка профессиональных кадров в магистратуре для цифровой экономики (ПКМ-2023) : Всероссийская научно-техническая и научно-методическая конференция магистрантов и их руководителей; Сборник лучших докладов: в 2 томах, Санкт-Петербург, 05–07 декабря 2023 года. Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2023. С. 328–334.*
27. Hore A., Ziou D. Image quality metrics: PSNR vs. SSIM // *2010 20th International Conference on Pattern Recognition. IEEE*, 2010. p. 2366–2369.
28. Sara U., Akter M., Uddin M. S. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study // *Journal of Computer and Communications*. 2019. v. 7. №. 3. p. 8–18.
29. Бардасов С. А. Эконометрика: учебное пособие. 2-е издание, переработанное и дополненное. // Тюмень: Издательство Тюменского государственного университета, 2010. 264 с.
30. Ченский А. А., Березкин А. А. Простые динамические модели предсказания времени работы вариационных автокодировщиков // *Материалы 80-й научно-технической конференции СПб НТО РЭС им. А. С. Попова, посвященной Дню радио, Санкт-Петербург, 21–25 апреля 2025. С.*
31. Roerich // Github. URL: <https://github.com/HSE-LAMBDA/roerich> (дата обращения: 05.01.2025).
32. Ikatser I. D., Kozitsin V. O. Skoltech anomaly benchmark (SKAB) // Kaggle. URL: <https://www.kaggle.com/dsv/1693952> (дата обращения: 05.01.2025).

33. *Shewhart W. A.* Economic Control of Manufactured Product. New York: D. Van Nostrand, Co, 1931. 501 с.

Березкин Александр Александрович

к.т.н., доцент кафедры программной инженерии и вычислительной техники (ПИиВТ), Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ, 193232, Санкт-Петербург, пр. Большевиков, д. 22, к. 1), директор Центра перспективных проектов и разработок, тел. +7 921 791-61-20, e-mail: berezkin.aa@sut.ru, ORCID ID: 0000-0002-1748-8642.

Ченский Александр Александрович

магистрант кафедры ПИиВТ, инженер Центра перспективных проектов и разработок, СПбГУТ, тел. +7 911 091-35-57, e-mail: chenski.aa@sut.ru, ORCID ID: 0009-0005-0832-8590.

Киричек Руслан Валентинович

д.т.н., ректор СПбГУТ, тел. +7 (812) 305-12-00, e-mail: kirichek@sut.ru, ORCID ID: 0000-0002-8781-6840.

Авторы прочитали и одобрили окончательный вариант рукописи.

Авторы заявляют об отсутствии конфликта интересов.

Вклад соавторов: *Каждый автор внёс свою долю участия как во все этапы проводимого теоретического исследования, так и при написании разделов данной статьи в следующих долях: Березкин А. А. – 70%, Ченский А. А. – 20%, Киричек Р. В. – 10%.*

Research of Video Stream Intensity Limits in UAV FPV Control in Frame Prediction Mode.

Part II: adaptive system of video stream latency and intensity control

Alexander A. Berezkin¹, Alexander A. Chenskiy¹, Ruslan V. Kirichek¹

¹ Bonch-Bruevich State University of Telecommunications (SPbSUT)

Abstract: In recent years, unmanned aircraft have been receiving active use. In a number of applications, first-person control of unmanned aircraft is used. First-person control requires the transmission of a video stream from an unmanned aircraft to an external pilot station. However, during network transmission, delay and frame loss of the video stream occurs. This results in lower intensity and higher latency of the video stream on the monitor of the external pilot station. In the previous part of the research, the method of increasing the intensity of video stream was presented, its mathematical models were built. In related works: the possibility of reducing the time of first-person control desynchronization was researched; the structure of unmanned systems first person control information exchange system was presented. The second part of the present research summarizes the results obtained in earlier works and presents an adaptive system of video stream latency and intensity control. The system consists from three blocks: neural codec mode controller, information exchange links control unit and video intensity control unit in two modes: intensity enhancement and delay reduction.

Keywords: neural network, variational autoencoder, first person view control, FPV, unmanned aerial system, unmanned aerial vehicle, video stream, prediction, prediction model, video stream prediction, DMVFN, application latency. video stream latency, adaptive control system.

For citation: Berezkin A. A., Chenskiy A. A., Kirichek R. V. Research of Video Stream Intensity Limits in UAV FPV Control in Frame Prediction Mode. Part II: adaptive system of video stream latency and intensity control. *Vestnik SibGUTI*, 2025, vol. 19, no. 3, pp. 139–162.
<https://doi.org/10.55648/1998-6920-2025-19-3-139-164>.



Content is available under the license
 Creative Commons Attribution 4.0
 License

© Berezkin A. A., Chenskiy A. A.,
 Kirichek R. V., 2025

The article was submitted: 13.03.2025;
 revised version: 28.04.2025;
 accepted for publication 28.04.2025.

References

1. Berezkin A. A., Chenskiy A. A., Kirichek R. V. Reduction of FPV video stream latency with future video stream frame prediction [Umen'shenie zaderzhki peredachi FPV-videopotoka cherez sputnikovye kanaly sv'azi metodom predskazaniya budushhih kadrov]. *Elektrosvjaz*. 2025. No. 4 (In Russ.)
2. Berezkin A. A., Chenskiy A. A., Kirichek R. V. Research of video stream intensity limits in UAV FPV control in frame prediction mode. Part I: models and methods [Issledovanie granic intensivnosti videopotoka pri FPV-upravlenii BPLA v rezhime predskazaniya kadrov. Chast' I: modeli i metody]. *Vestnik SibGUTI*, 2024, No. 18(3), pp. 115-139.
3. Hu X., Huang Z., Huang A., Xu J., Zhou S. A dynamic multi-scale voxel flow network for video prediction. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6121-6131.
4. Choquette J., Gandhi W. Nvidia A100 GPU: Performance & innovation for GPU computing. *2020 IEEE Hot Chips 32 Symposium (HCS), IEEE Computer Society*, 2020, pp. 1-43.
5. Choquette J., Gandhi W., Giroux O., Stam N., Krashinsky R. Nvidia A100 tensor core GPU: Performance and innovation. *IEEE Micro*, 2021, No. 41(2), pp. 29-35.
6. Giambene G., Kota S., Pillai P. Satellite-5G integration: A network perspective. *IEEE Network*, 2018, No. 32(5), pp. 25-31.
7. Berezkin A. A., Parshin A. A., Lazarev A. A. Adaptivnyj kontrol' intensivnosti videopotoka pri peredache FPV-trafika bespilotnyh sistem [Adaptive control of video stream intensity in FPV traffic transmission of unmanned systems]. *79-ja nauchno-tehnicheskaja konferencija SPb NTO RES im. A.S. Popova, posvjashhennaja Dnju radio: sbornik dokladov regional'noj konferencii*, Saint Petersburg, 22-26 April, 2024, pp. 162-165.
8. Strategy for the development of the telecommunications industry of the Russian Federation for the period until 2035 [Strategija razvitiya otrasli svjazi Rossijskoj Federacii na period do 2035 goda]. *Government of Russia*. URL: [texttthttp://government.ru/news/50304/](http://government.ru/news/50304/) (accessed: 02.01.2025).
9. Berezkin A. A., Chenskiy A. A., Kirichek R. V., Zaharov A. A. Research of neural codecs configurations for adaptive unmanned systems FPV-control video stream frame compression system. Part I. Methodics [Issledovanie konfiguracij nejrosetevyh kodekov dlja adaptivnoj sistemy szhatija kadrov FPV-videopotoka pri upravlenii bespilotnymi sistemami. Chast' I. Metodika]. *Elektrosvjaz*, Sankt-Peterburg, 2024, No. 9, pp. 42-51.
10. Berezkin A. A., Chenskiy A. A., Kirichek R. V., Zaharov A. A. Research of neural codecs configurations for adaptive unmanned systems FPV-control video stream frame compression system. Part II. Experiment [Issledovanie konfiguracij nejrosetevyh kodekov dlja adaptivnoj sistemy szhatija kadrov FPV-videopotoka pri upravlenii bespilotnymi sistemami. Chast' II. Jeksperiment]. *Elektrosvjaz*, Sankt-Peterburg, 2024, No. 10, pp. 59-69.
11. Zhou D., Wang W., Yan H., Lv W., Zhu Y., Feng J. MagicVideo: Efficient video generation with latent diffusion models. *arXiv preprint arXiv:2211.11018*, 2022.
12. Saharia C., Ho J., Chan W., Salimans T., Fleet D. J., Norouzi M. Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence*, 2022, No. 45(4), pp. 4713-

4726.

13. Liu Z., Lin Y., Cao Y., Hu H., Wei Y., Zhang Z., Lin S., Guo B. SWIN transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10012-10022.
14. Liang J., Cao J., Sun G., Zhang K., Van Gool L., Timofte R. SwinIR: Image restoration using swin transformer. *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 1833-1844.
15. Wang X., Yu K., Wu S., Gu J., Liu Y., Dong C., Qiao Y., Loy C. C. ESRGAN: Enhanced super-resolution generative adversarial networks. *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018.
16. Kozlova A. I., Oblakov N. A., Berezkin A. A. Comparative analysis of superresolution methods in UAV FPV-control video stream compression system [Sravnitel'nyj analiz metodov sverhrazreshenija v sisteme szhatija videopotoka pri FPV upravlenii bespilotnymi sistemami] // *Materialy 79-j nauchno-tehnicheskoy konferencii SPb NTO RJeS im. A. S. Popova, posvjashhennoj Dnju radio*, Saint Petersburg, 22–26 April 2024, pp. 166–169.
17. Berezkin A. A., Vivchar' R. M., Slepnev A. V., Kirichek R. V., Zaharov A. A. Video stream compression method for unmanned systems first person view controll in hybrid orbital-terrestrial networks [Metod szhatija videopotoka pri upravlenii bespilotnymi sistemami v gibridnyh orbital'no-nazemnyh setjah svjazi]. *Jelektrosvjaz'*, 2023, №. 10, pp. 48-56.
18. Berezkin A. A., Vivchar' R. M., Kirichek R. V., Zaharov A. A. Method of FPV-video stream from unmanned systems decompression with diffusional neural model [Metod dekompressii FPV-videopotoka ot bespilotnyh sistem na osnove latentnoj diffuzionnoj nejrosetevoj modeli]. *Elektrosvjaz'*, 2024, №. 1, pp. 42-53.
19. Berezkin A. A., Chenskiy A. A., Kirichek R. V., Zaharov A. A. Masking of UAV FPV-control video stream. Part I: Models and methodics [Maskirovanie kodirovannogo FPV videopotoka pri upravlenii BPLA. Chast' I. Modeli i metodika]. *Elektrosvjaz'*, 2024, №. 12-2, pp. 52-65.
20. Berezkin A. A., Chenskiy A. A., Kirichek R. V. Video stream frames artifacts suppression in unmanned systems FPV-control [Nivelirovanie artefaktov kadrov videopotoka pri FPV-upravlenii bespilotnymi sistemami]. *Infkommunikacionnye tehnologii*, 2024, № 22(85), pp. 7-17.
21. Berezkin A. A., Oblakov N. A., Chenskiy A. A., Kirichek R. V. Latent video stream correlational analysis in UAV first person view control [Korreljacionnyj analiz latentnogo videopotoka pri upravlenii BAS ot pervogo lica]. *Elektrosvjaz'*, 2024, № 11, pp. 25-31.
22. Larzon L. A., Degermark M., Pink S. UDP lite for real time multimedia applications. *Hewlett-Packard Laboratories*, 1999.
23. Allman M., Falk A. On the effective evaluation of TCP. *ACM SIGCOMM Computer Communication Review*, 1999, №. 29(5), pp. 59-70.
24. Langley A., Riddoch A., Wilk A., Vicente A., Krasic C., Zhang D., Shi Z. The QUIC transport protocol: Design and internet-scale deployment. *Proceedings of the conference of the ACM special interest group on data communication*, 2017, pp. 183-196.
25. Cui Y., Li T., Liu C., Wang X., Kühlewind M. Innovating transport with QUIC: Design approaches and research challenges. *IEEE Internet Computing*, 2017, №. 21(2), pp. 72-76.
26. Berezkin A. A., Matveev D. V., Savelov D. Ju. Video prediction model optimization with Tensorrt to decrease latency in UAV control [Optimizacija modeli predskazanija video s pomoshh'ju tensorrt dlja umen'shenija zaderzhek pri upravlenii BPLA]. *Podgotovka professional'nyh kadrov v magistrature dlja cifrovoj jekonomiki (PKM-2023) : Vserossijskaja nauchno-tehnicheskaja i nauchno-metodicheskaja konferencija magistrantov i ih rukovoditelej; Sbornik luchshih dokladov: v 2 tomah*, Saint Petersburg, 05–07 december 2023, Saint Petersburg: Bonch-Bruevich State University of Telecommunications, 2023, pp. 328-334.
27. Hore A., Ziou D. Image quality metrics: PSNR vs. SSIM. *2010 20th international conference on pattern recognition*, IEEE, 2010, pp. 2366-2369.
28. Sara U., Akter M., Uddin M. S. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. *Journal of Computer and Communications*, 2019, No. 7(3), pp. 8-18.

29. Bardasov S. A. Econometrics: textbook. 2nd edition, revised and expanded [Jekonometrika: uchebnoe posobie. 2-e izdanie, pererabotannoe i dopolnennoe]. Tjumen': Izdatel'stvo Tjumenskogo gosudarstvennogo universiteta, 2010, 264 P.
30. Chenskiy A. A., Berezkin A. A. Simple dynamic models for variational autoencoders inference time prediction [Prostye dinamicheskie modeli predskazaniya vremeni raboty variacionnyh avtokodirovshhikov] // *Materialy 80-j nauchno-tehnicheskoy konferencii SPb NTO RJeS im. A. S. Popova, posvjashhennoj Dnju radio*, Saint Petersburg, 21–25 April 2025, pp.
31. Roerich. *Github*. URL: <https://github.com/HSE-LAMBDA/roerich> (accessed: 05.01.2025).
32. Ikatser I. D., Kozitsin V. O. Skoltech anomaly benchmark (SKAB). *Kaggle*. URL: <https://www.kaggle.com/dsv/1693952> (accessed: 05.01.2025).
33. Shewhart W. A. Economic Control of Manufactured Product. *New York: D. Van Norsrand, Co*, 1931, 501 P.

Berezkin Alexander Alexandrovich

PhD, associate professor of the Software Engineering and Computer Science Department, Bonch-Bruевич State university of telecommunications (SPbSUT, 193232, Saint-Petersburg, Bolshevnikov Avenue, 22-1), phone. +7 921 791-61-20, e-mail: berezkin.aa@sut.ru, ORCID ID: 0000-0002-1748-8642.

Chenskiy Alexander Alexandrovich

Master grade student of the Program Engineering and Computer Science Department, engineer of Center of Perspective Projects and Developments, SPbSUT, phone. +7 911 091-35-57, e-mail: chenskii.aa@sut.ru, ORCID ID: 0009-0005-0832-8590..

Kirichek Ruslan Valentinovich

Doctor of technical science, rector of SPbSUT, phone. +7 (812) 305-12-00, e-mail: kirichek@sut.ru, ORCID ID: 0000-0002-8781-6840.