

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ТЕХНОЛОГИЙ ПЕРЕДАЧИ И ОБРАБОТКИ ИНФОРМАЦИИ И СИГНАЛОВ

УДК 004.622

НИВЕЛИРОВАНИЕ АРТЕФАКТОВ КАДРОВ ВИДЕОПОТОКА ПРИ FPV-УПРАВЛЕНИИ БЕСПИЛОТНЫМИ СИСТЕМАМИ

*Березкин А.А., Ченский А.А., Киричек Р.В.**Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, РФ**E-mail: berezkin.aa@sut.ru, chenskii.aa@sut.ru, kirichek@sut.ru*

При использовании пакетных сетей для управления беспилотными системами от первого лица возникает проблема большого объема передаваемых данных. Наибольший объем трафика при управлении от первого лица формируют кадры видеопотока. Соответственно, для повышения эффективности использования сети связи между беспилотными системами и станцией внешнего пилота необходимо осуществлять сжатие кадров видеопотока. Высокую степень сжатия обеспечивает использование вариационных автокодировщиков. Одной из проблем использования вариационных автокодировщиков для сжатия изображений является возникновение на изображениях специфичных артефактов. В настоящей статье предлагаются способы нивелирования возникновения артефактов при восстановлении изображений из латентного пространства нейросетевым декодером, а также предлагается эмпирическая шкала оценки артефактов автокодировщиков. Предложенный подход заключается в подготовке пиксельных данных кадра видеопотока к кодированию и восстановлению их после декодирования. Экспериментальным путем показано, что один из предложенных методов позволяет устранять абсолютное большинство артефактов без внесения существенных искажений в восстанавливаемые кадры.

Ключевые слова: *нейронная сеть, подавление артефактов, нивелирование артефактов, артефакты изображения, передача видеопотока, вариационный автокодировщик, нейросетевой кодек, FPV-управление, управление от первого лица, беспилотные системы, беспилотные воздушные суда*

Введение

Беспилотные системы (БС), такие как беспилотные авиационные системы (БАС), на текущий момент широко распространены и применяются в целом ряде различных областей [1–2]. Широко применяемым способом управления БС является управление от первого лица (FPV-управление). В ходе FPV-управления существует две основные линии связи между БС и станцией внешнего пилота (СВП): линия связи с землей и линия связи с бортом. Первая используется для передачи данных с БС на СВП. К передаваемым данным относятся: идентификатор, координаты, скорость, курсовой угол, высота, метаданные, кадры видеопотока. Вторая используется для передачи данных с СВП на БС. По ней передаются команды управления БС.

Линии связи реализуются посредством радиосигналов либо с использованием существующих сотовых сетей передачи данных, которые имеют ограниченную зону покрытия и пропускную способность. Перспективным направлением для обеспечения линий связи при FPV-управлении является использование гибридных орбитально-наземных сетей связи [1–3]. Тем не менее, задействование космического сегмента при передаче данных в сетях связи приводит к существен-

ному увеличению задержки, джиттера и уменьшению пропускной способности в каналах связи.

Это приводит к необходимости снижения объема трафика при FPV-управлении. Наибольший объем при этом приходится на трафик кадров видеопотока. Соответственно, сжатие видеопотока является ключевым фактором эффективного использования пакетных сетей при FPV-управлении. Ограниченная пропускная способность космического сегмента сетей связи вынуждает искать решения по обеспечению степени сжатия, превышающей предлагаемую с помощью стандартных средств сжатия кадров видеопотока и отдельных изображений с потерями, к которым относятся: кодеки h264 [4], h265 [5], форматы изображений JPEG [6], AVIF [7], WebP [8].

Намного более высокой степени сжатия кадров видеопотока позволяют добиться вариационные автокодировщики [1–2; 9]. Тем не менее, методы, в которых они используются, характеризуются увеличением искажения восстанавливаемых на стороне СВП кадров видеопотока. Кроме того, появляются специфичные для вариационных автокодировщиков артефакты восстанавливаемых изображений.

Артефакты изображений представляют собой эффекты и объекты на вторичных изображениях,

которые не присутствуют на первичных. Они возникают при сжатии изображений с потерями, наличии быстрых движений в видеопотоке и различных световых эффектах при съемке. Первичные изображения являются источниками для вторичных изображений. В случае сжатия изображений – это изображение до сжатия, в случае фото- и видеосъемки – это представление реальности.

К видам артефактов изображений относят: артефакты сжатия, появляющиеся при достаточно сильном сжатии с потерями (широко известны, в частности, артефакты сжатия алгоритмом JPEG [10]); хроматические aberrации, при которых по краям объектов создаются цветовые контуры; муаровые узоры, возникающие при наложении сетчатых рисунков [11]; артефакты движения, возникающие при съемке движущихся объектов [12].

Использование вариационных автокодировщиков приводит к появлению иного вида артефактов: артефактов автокодировщиков (рисунок 1). Они представляют собой скопления пикселей белого, синего, голубого, желтого и розового цветов зачастую округлой формы и часто проявляются в темных и светлых областях изображений. Тем не менее, их распространение также охватывают окружающие группы пикселей. Иногда данные пиксели бывают рассеяны.

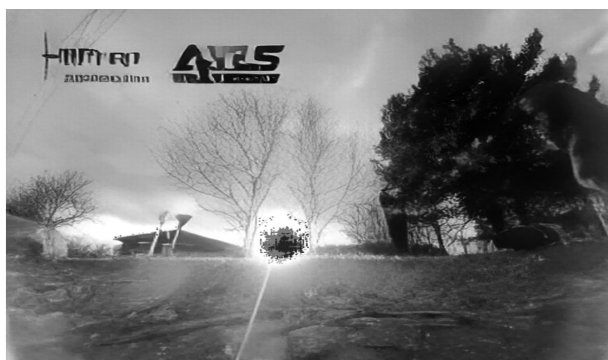


Рисунок 1. Пример артефактов автокодировщика VQ-f16 без квантования

Алгоритмы квантования и сжатия латентного пространства оказывают влияние на проявляемые артефакты (рисунок 2). Тем не менее, артефакты автокодировщиков проявляются и при отсутствии воздействия на латентное пространство, полученное из кодируемых изображений.

В некоторых моделях вариационных автокодировщиков артефакты более выражены, чем в других (рисунок 3).

В отличие от хроматических aberrаций, не слишком сильных артефактов сжатия и муаровых узоров, артефакты автокодировщиков, когда они проявляются в сплошной, а не в рассеянной

форме, не позволяют воспринимать признаковые структуры тех частей изображения, которые за ними скрыты. Данные артефакты изображений серьезно усложняют восприятие внешним пилотом (ВП) полетной обстановки и, таким образом, приводят к возникновению рисков провала полетного задания и потери БС при FPV-управлении.



Рисунок 2. Пример артефактов автокодировщика VQ-f16 при степенном квантовании



Рисунок 3. Пример артефактов автокодировщика KL-f16 при линейном квантовании

Для нивелирования артефактов изображений используются как алгоритмические методы [13; 14], так и нейронные сети [15–17]. Тем не менее, использование нейронных сетей при нивелировании артефактов автокодировщиков нежелательно. Во-первых, это неизбежно приведет к искажению исходного изображения вследствие существенной степени сокрытия исходных признаковых структур артефактами. Во-вторых, нейронные сети, как правило, требуют больше времени на выполнение, чем обычные алгоритмы, что приводит к уменьшению частоты кадров в секунду (FPS) на стороне ВП. Это негативно сказывается на возможности оперативно реагировать на изменяющуюся полетную обстановку и дистанционно пилотировать БС. В-третьих, в данном случае есть доступ к исходному изображению (на стороне БС) и информации об областях концентрации артефактов.

Исходя из вышеизложенного, целесообразнее будет осуществлять алгоритмическую подготовку пиксельных данных изображения перед кодированием и опциональное восстановление пиксельных данных после декодирования таким образом, чтобы области возникновения артефактов автокодировщиков нивелировались.

Кроме того, необходима методика для оценки работоспособности алгоритмов нивелирования артефактов автокодировщиков. Для оценки качества восстанавливаемого изображения по сравнению с исходным широко используются метрики качества SSIM, MSE и PSNR [1–2; 9]. Тем не менее, данные метрики оценивают качество всего восстанавливаемого изображения целиком, а не отдельных его элементов. Предпочтительным является использование некоторой эмпирической шкалы, которая бы позволяла оценивать наличие артефактов именно с точки зрения визуального восприятия восстанавливаемого изображения ВП.

В настоящей работе предлагаются алгоритмы нивелирования артефактов вариационных автокодировщиков и проводится их эксперимен-

тальная оценка по предлагаемой методике, использующей эмпирическую шкалу артефактов автокодировщиков.

Нейросетевой кодек

Предлагаемые в данной работе алгоритмы нивелирования артефактов предназначены для использования в рамках нейросетевого кодека передачи FPV-видеопотока. Их возможные области применения, тем не менее, не ограничиваются нейросетевыми кодеками и FPV-управлением.

В данной работе используется нейросетевой кодек, основанный на представленном в исследовании [1]. Он состоит из кодера (рисунок 4) и декодера (рисунок 5). Нейросетевой кодек размещается на БС и предназначен для осуществления процессов сжатия кадров. Нейросетевой декодер размещается на СВП и предназначен для восстановления кадров с целью дальнейшего их отображения ВП.

f – пиксельные данные кадра в разрешении 1280×720 до интерполяции. f представляет собой многомерный массив типа данных uint8 размер-

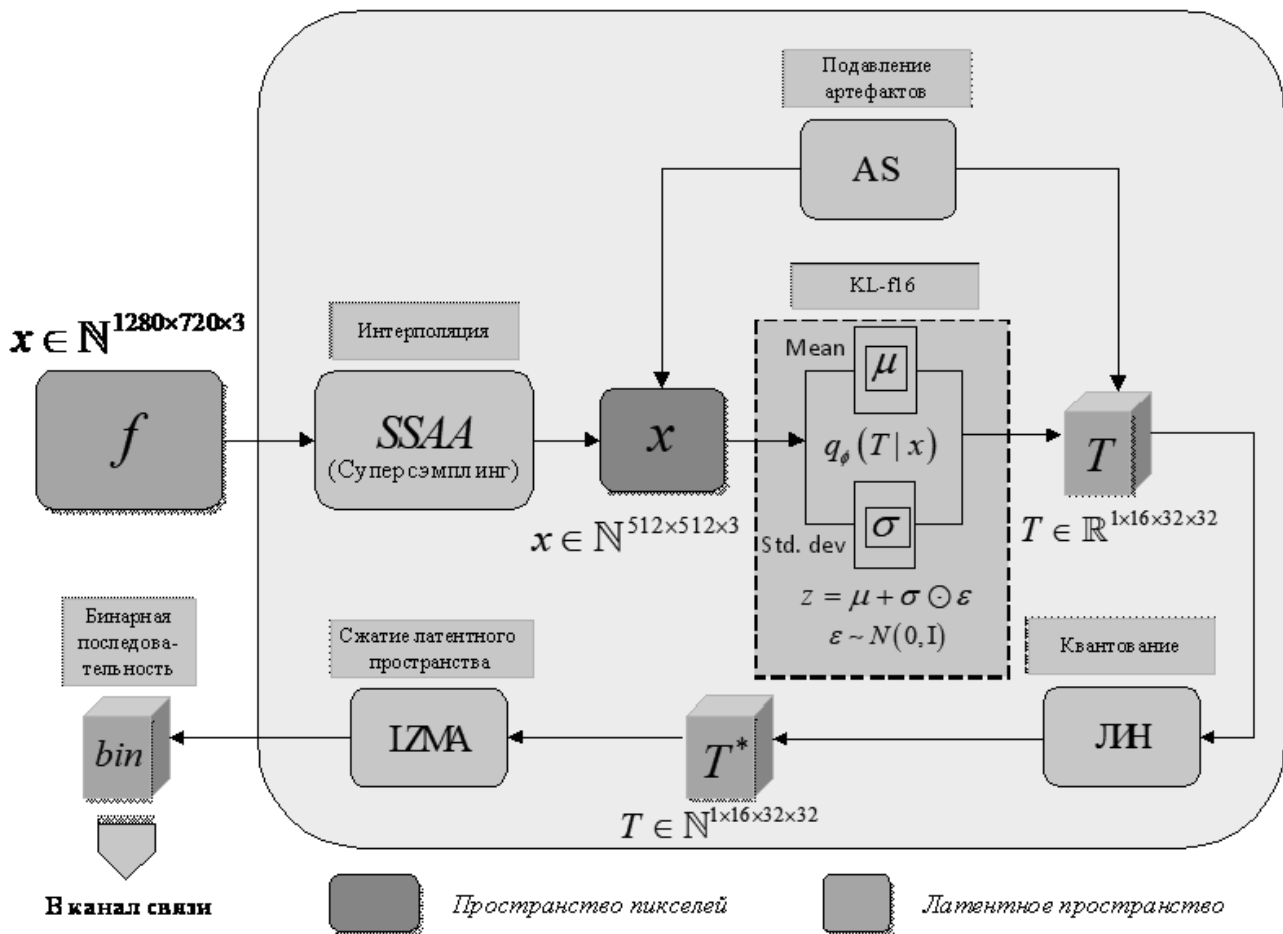


Рисунок 4. Нейросетевой кодек

ности (720, 1280, 3), где 3 – число каналов (RGB). Значения пикселей: от 0 до 255.

x – пиксельные данные интерполированного кадра в разрешении 512×512. В отличие от f , x переводятся в формат тензора типа данных float16. Размерность: (3, 512, 512). Значения пикселей: от 0 до 1 в соответствии с формулой:

$$\forall i, j, k : \begin{cases} i \in [0; 511], i \in N \cup \{0\} \\ j \in [0; 511], j \in N \cup \{0\} : x_{ijk} = \frac{x_{ijk}^*}{255}, \\ k \in \{0, 1, 2\} \end{cases}$$

где i, j и k – индексы значения тензора, x_{ijk} – значение тензора x , а x^* – соответствующие uint8 значения.

T – тензор латентного пространства типа данных float16 размерности (1, 8, 32, 32).

T^* – квантованный к типу данных uint8 тензор T .

bin – сжатая бинарная последовательность для передачи по линии связи с землей.

Нейросетевой кодек для исследования [1] предполагает возможность использования нескольких алгоритмов сжатия. В данном исследовании используется его частный случай с алгоритмом сжатия LZMA.

В качестве вариационного автокодировщика вместо VQ-f16 из исследования [1] используется KL-f16 из моделей Stable Diffusion [18]. Причи-

ной этого является большая выраженность артефактов автокодировщиков в последнем.

Параметры линейного квантования (ЛИН), рассчитаны для KL-f16 и равны -20,96 и 7,78.

Алгоритмы нивелирования артефактов автокодировщиков

Предлагаемые алгоритмы нивелирования артефактов основываются на гипотезе о том, что артефакты автокодировщиков часто проявляются в темных и светлых областях. В этой связи источник их возникновения – группы близких по RGB-значениям пикселей, значения R, G и B которых вместе близки к 0 или 255. Таким образом, кодирование данных областей приводит к сильному отклонению значений латентного пространства, что, в свою очередь, создает проблему восстановления данных областей при декодировании.

Алгоритмы нивелирования артефактов автокодировщиков могут применяться на трех этапах: предобработке, промежуточной обработке и постобработке. Предварительная обработка выполняется на стороне нейросетевого кодера до кодирования. Промежуточная обработка выполняется после кодирования и до этапа декодирования, затрагивая латентное пространство, и может выполняться как на кодере, так и на де-

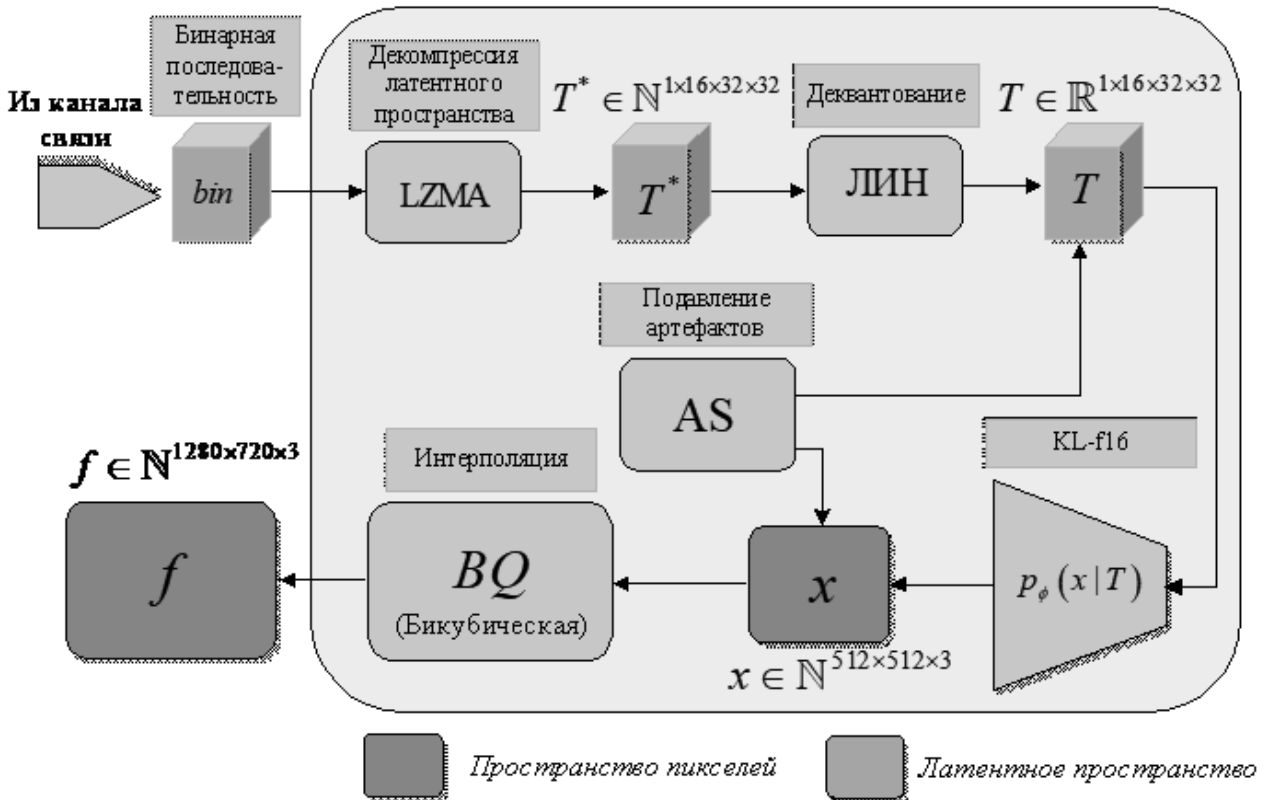


Рисунок 5. Нейросетевой декодер

кодере. Постобработка выполняется после декодирования на восстановленном изображении на стороне СВП. При этом, как предобработка, так и постобработка могут выполняться как с исходными пиксельными данными в формате пикселей uint8 (стандартный вариант RGB – f), так и в формате интенсивности отдельных цветов пикселей в float16 (x). Промежуточная обработка может затрагивать как квантованный тензор T^* , так и не квантованный T . В рамках одного алгоритма нивелирования артефактов может выполняться несколько видов обработки.

Первый алгоритм нивелирования артефактов ASMoveDistribution (перенос распределения) не является новой разработкой и был обнаружен путем проведения анализа кода программных реализаций таких систем сжатия изображений, как Stable Diffusion [18] и CDC Compression (Lossy Image Compression with Conditional Diffusion Models) [19]. В рамках данного алгоритма осуществляется как предобработка, так и постобработка, затрагивающие x . Пусть x' – новые значения тензора, тогда при предобработке они рассчитываются в соответствии с выражением:

$$x' = x \times 2 - 1,$$

а при постобработке исходные значения восстанавливаются в соответствии с выражением:

$$x = \begin{cases} 1, \frac{x'+1}{2} > 1 \\ \frac{x'+1}{2}, \frac{x'+1}{2} \in [0;1] \\ 0, \frac{x'+1}{2} < 0 \end{cases}.$$

Второй алгоритм нивелирования артефактов ASCutEdgeColors (отсечение крайних цветов), одна из новых разработок, которая предлагается в настоящей работе. В отличие от ASMoveDistribution, предлагаемый метод является параметрическим с входным параметром delta. Используется только в предобработке применительно к f . Его принцип работы заключается в замене значений пикселей, значения всех RGB-цветов которых близки к 255 ближе delta и ближе, на (255 – delta):

$$\forall i, j : f'_{ij} = \begin{cases} f_{ij}, \exists k \in \{0,1,2\} : f_{ijk} \notin (255 - \delta, 255] \\ [255 - \delta, 255 - \delta, 255 - \delta], \\ k \in \{0,1,2\} : f_{ijk} \in (255 - \delta, 255] \end{cases},$$

где f' – новые значения f .

Таким образом, убираются области пикселей, которые, согласно гипотезе, являются источниками данных артефактов. При достаточно малых

значениях delta изменения на изображении незаметны для ВП и не оказывают влияния на возможность FPV-управления БС.

Третий алгоритм нивелирования артефактов ASComposit (композит), является комбинацией первых двух алгоритмов: переноса распределения и отсечения крайних цветов.

Четвертый алгоритм нивелирования артефактов ASCutEdgeValues (отсечение крайних значений) предназначен для промежуточной обработки тензора латентного пространства T и T^* . Он основывается на гипотезе о том, что области латентного пространства, которые при декодировании обращаются в артефакты изображения, являются значениями латентного пространства, достаточно отдаленными от среднего. Как и первый и третий алгоритмы, он также является параметрическим, принимая на вход значение K – коэффициент межквартильного интервала IQR. Новые значения T' получаются по формуле:

$$\forall i : T'_i = \begin{cases} Q_1, T_i < Q_1 - K \times IQR \\ T_i, T_i \in [Q_1 - K \times IQR; Q_3 + K \times IQR], \\ Q_3, T_i > Q_3 + K \times IQR \end{cases},$$

где Q_1 – первый квартиль, Q_3 – второй квартиль, а i – наборы индексов тензора латентного пространства T .

Его главная особенность заключается в возможности использования при обработке латентного пространства без информации об исходном изображении.

Пятый алгоритм нивелирования артефактов ASLatentComposit (латентный композит) является комбинацией первого (ASMoveDistribution) и четвертого (ASCutEdgeValues) алгоритмов. Он принимает на вход значение K .

Эмпирическая шкала оценки артефактов автокодировщиков

Для оценки методов нивелирования артефактов автокодировщиков необходима шкала их оценки. В настоящей работе предлагается пяти- (шести-) бальная шкала (таблица 1). В случае, когда наличие артефактов подразумевается изначально и есть необходимость их оценить, следует использовать сокращенный (пятибальный) вариант шкалы: без значения 0 – чистого изображения. В противном случае – ее полный (шестибальный) вариант.

Ключевым критерием оценки по данной шкале является их влияние на возможность восприятия исходного изображения. Влиянию воз-

возможности восприятия соответствует характер (сплошной либо распределенный), расположение (локализация в одном или нескольких местах) и масштаб (малый, средний, большой) артефактов автокодировщиков.

Таблица 1. Эмпирическая шкала артефактов автокодировщиков

#	Название	Описание
0	Чистое изображение	Артефакты автокодировщиков отсутствуют
1	Пиксельная рябь	Артефакты автокодировщиков представляют собой распределенные пиксели в темных или светлых областях. Не создают помех для восприятия исходного изображения
2	Малые сплошные артефакты	На изображении присутствует одна-две области сплошных артефактов автокодировщиков. Сами области составляют незначительную часть от всего изображения. Помех для восприятия исходного изображения кроме данных незначительных областей нет
3	Средние сплошные артефакты	На изображении присутствуют три и более области сплошных артефактов малых и средних размеров. Присутствуют помехи для восприятия ряда областей изображения, но восприятие изображения в целом возможно
4	Большие сплошные артефакты	Наличие одной или нескольких областей сплошных артефактов автокодировщиков, покрывающих значительную часть изображения. Существенные помехи в восприятии изображения. Есть возможность восприятия лишь отдельных областей, имеются затруднения с пониманием общего содержания изображения
5	Полные сплошные артефакты	Сплошные артефакты покрывают все изображение полностью или абсолютно большую его часть. Восприятие исходного изображения невозможно

В случае признаков нескольких оценок следует выбирать наибольшую оценку. Например,

при наличии двух малых областей сплошных артефактов и пиксельной ряби в нескольких других местах следует выбирать оценку 2 (малые сплошные артефакты), не учитывая пиксельную рябь.

Предлагаемая шкала является эмпирической. Соответственно, возможны пограничные случаи, когда возможно поставить обе оценки. В таких случаях рекомендуется оценить, насколько велика помеха от артефактов для восприятия изображения.

Набор данных

В настоящем исследовании используется авторский набор данных, состоящий из 100 кадров, извлеченных из видеозаписи полета БАС (рисунок 6). Кадры данной видеозаписи характеризуются наличием резко светлых и темных областей, что приводит к высокой частоте появления артефактов автокодировщиков.



Рисунок 6. Примеры кадров набора данных

Извлечение проведено равномерно из всей видеозаписи согласно методике, разработанной в исследовании [1] и также применявшейся в работе [2].

Методика экспериментов

В рамках одного эксперимента выполняется преобразование всех 100 кадров набора данных нейросетевым кодеком: кодером и декодером последовательно. Затем сто кадров оцениваются по предлагаемой эмпирической шкале артефактов автокодировщиков.

Конфигурации экспериментов:

1. Без нивелирования артефактов.
2. Перенос распределения.

3. Отсечение крайних цветов. Значения параметра δ : 5, 15, 20, 25, 30.

4. Композит. Значения параметра δ : 5, 15, 20, 25, 30.

5. Отсечение крайних значений. Значения параметра K : 0,50, 1,00, 1,50.

6. Латентный композит при $K = 1,50$.

Итого проводится 16 экспериментов.

Эксперименты

В результате проведения экспериментов была получена оценка количества кадров при каждой конфигурации в соответствии с каждым значением эмпирической шкалы артефактов автокодировщиков (таблица 2).

Визуальные искажения при применении всех рассматриваемых видов нивелирования артефактов являются незначительными, но в отдельных случаях присутствуют (рисунок 7). При значениях δ более 15 в ASCutEdgeColors и ASComposit проявляется искажение наиболее светлых и темных областей, что становится заметно при просмотре кадров. В случае ASCutEdgeValues при $K < 1,5$ проявляются заметные искажения в визуальных чертах изображений.

При ASMoveDistribution сплошные артефакты частично сохраняются, но в большинстве случаев становятся очень малозаметными. Кроме того, полностью исчезает пиксельная рябь.

ASCutEdgeColors эффективно устраняет большинство артефактов, но отдельные из них (которые не встречались в VQ-f16, но встречаются в KL-f16) не удается устранить с помощью данного метода.

Метод ASComposit позволяет полностью устранить пиксельную рябь и большую часть артефактов, а сохраняющиеся артефакты в подавляющем большинстве случаев являются малозаметными.

ASCutEdgeValues при $K \geq 1,00$ устраняет большинство артефактов, сохраняющиеся артефакты выглядят смазано и являются малозаметными, но метод малоприменим по причине существенных визуальных искажений изображения.

Лучше всего показывает себя алгоритм ASLatentComposit при $K = 1,50$. Он устраняет абсолютное большинство артефактов с небольшими цветовыми искажениями. Сохраняющиеся же артефакты смазаны и незаметны.

Полученные результаты показывают, что большинство артефактов автокодировщиков появляются из-за присутствия в латентном пространстве значений, достаточно далеких от их среднего значения. С другой стороны, не все артефакты автокодировщиков появляются в наиболее светлых и темных областях пиксельного пространства исходного изображения.

Заключение

В настоящем исследовании рассмотрены артефакты вариационных автокодировщиков и пред-

Таблица 2. Результаты экспериментов, проведенных с применением методов нивелирования артефактов

Номер эксперимента	Метод нивелирования артефактов	0	1	2	3	4	5
1	Без нивелирования артефактов	4	4	14	58	19	1
2	Перенос распределения	41	0	49	10	0	0
3	Отсечение крайних цветов: $\delta = 5$	4	12	7	57	19	1
4	Отсечение крайних цветов: $\delta = 10$	4	15	30	40	11	0
5	Отсечение крайних цветов: $\delta = 15$	5	13	35	40	7	0
6	Отсечение крайних цветов: $\delta = 20$	5	14	38	40	3	0
7	Отсечение крайних цветов: $\delta = 25$	5	14	38	40	3	0
8	Композит: $\delta = 5$	74	0	22	4	0	0
9	Композит: $\delta = 10$	78	0	18	4	0	0
10	Композит: $\delta = 15$	82	0	14	4	0	0
11	Композит: $\delta = 20$	80	0	16	4	0	0
12	Композит: $\delta = 25$	78	0	18	4	0	0
13	Отсечение крайних значений: $K = 0,50$	21	79	0	0	0	0
14	Отсечение крайних значений: $K = 1,00$	93	6	1	0	0	0
15	Отсечение крайних значений: $K = 1,50$	87	11	2	0	0	0
16	Латентный композит: $K = 1,50$	98	0	2	0	0	0

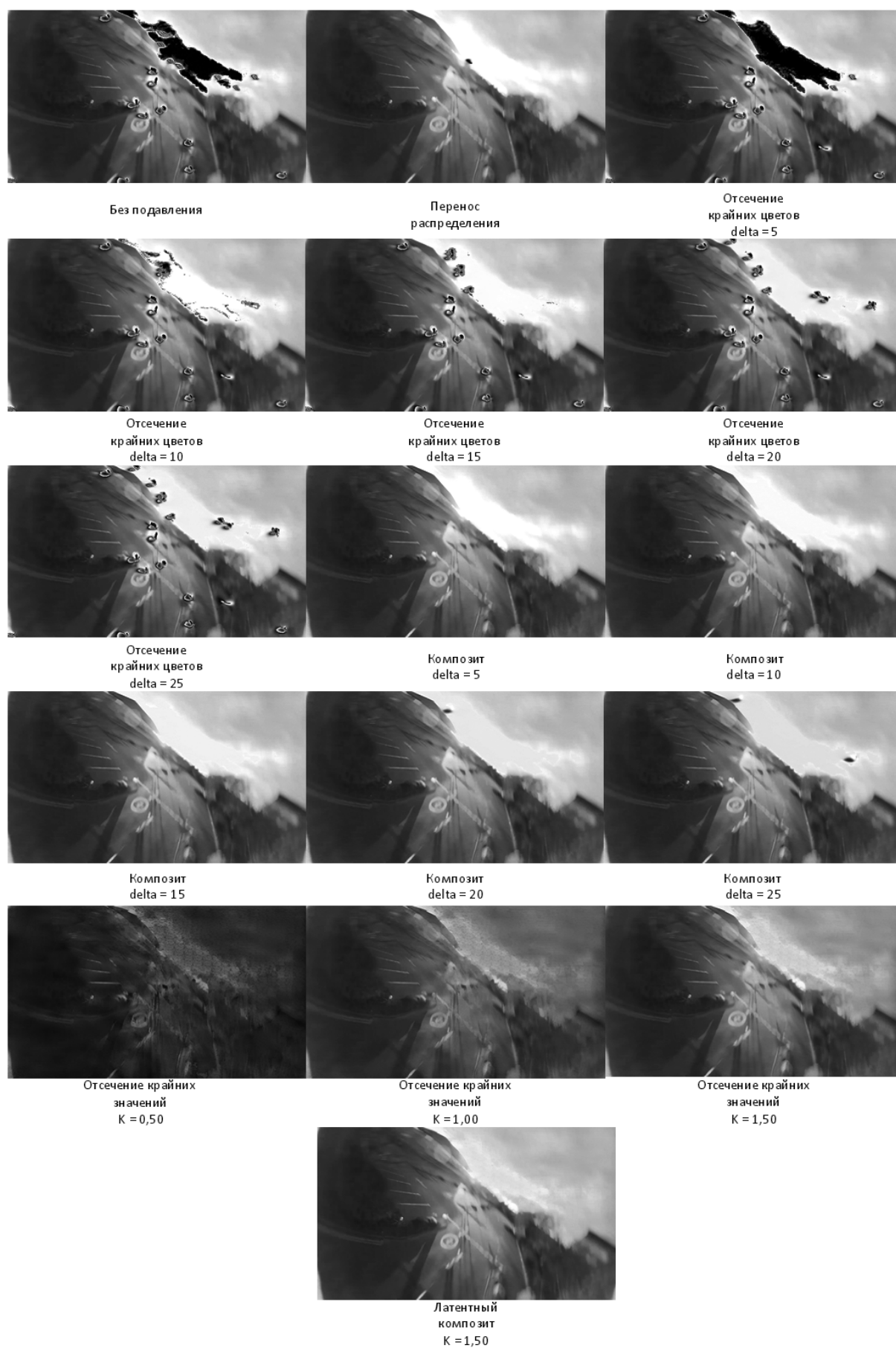


Рисунок 7. Эффект нивелирования артефактов с применением различных алгоритмов на примере одного кадра

ложена эмпирическая шкала их оценки в полном (шестибальная) и сокращенном (пятибальная) вариантах.

Рассмотрено пять алгоритмов нивелирования артефактов, четыре из которых являются новыми и один не требует вмешательства в исходное изображение.

Наибольшее количество артефактов позволяют убрать алгоритмы ASComposit с параметром $\delta = 15$ (81,25%), ASCutEdgeValues с параметром $K = 1,50$ (86,46%) и ASLatentComposit с параметром $K = 1,50$ (97,92%). Наименьшие искажения изображения из них наблюдаются у алгоритма ASComposit с параметром $\delta = 15$ (81,25%), который и рекомендуется использовать для нивелирования артефактов автокодировщиков.

С помощью алгоритма ASCutEdgeValues возможно нивелирование артефактов только на стороне нейросетевого декодера, для остальных алгоритмов требуются также манипуляции на стороне нейросетевого кодера.

Использование рассмотренных в настоящей работе алгоритмов позволяют устранять большинство артефактов автокодировщиков с небольшими визуальными искажениями. Данное исследование, таким образом, позволяет существенно облегчить решение проблемы появления артефактов на кадрах FPV-видеопотока при использовании нейросетевых автокодировщиков.

Научная работа выполнена в рамках прикладных научных исследований Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), регистрационный номер 1023031600087-9-2.2.4;2.2.5;2.2.6;1.2.1;2.2.3.

Литература

1. Исследование методов латентного сжатия видеопотока при FPV управлении беспилотными системами / А.А. Березкин [и др.] // Электросвязь. 2024. № 6. С. 26–36.
2. Исследование методов квантования латентного пространства вариационного автокодировщика для кадров FPV видеопотока. Часть I / А.А. Березкин [и др.] // Электросвязь. 2024. № 6. С. 10–16.
3. Проект стратегии развития отрасли связи Российской Федерации на период до 2035 года [утвержден Правительством Российской Федерации 24 ноября 2023 г.]. URL: <https://digital.gov.ru/ru/documents/9120/> (дата обращения: 27.07.2024).
4. Рекомендация МСЭ Т Н.264. Усовершенствованное кодирование изображения для общих аудиовизуальных услуг. Женева, 2008. 342 с.
5. Recommendation ITU-T H.265 (V9). High Efficiency Video Coding. Geneva, 2023. 718 p.
6. ISO/IEC 10918-1:1994. Information Technology – Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines. URL: <https://www.iso.org/ru/standard/18902.html> (дата обращения: 27.07.2024).
7. Barman N., Martini M.G. An evaluation of the next-generation image coding standard AVIF // 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX). 2020. P. 1–4.
8. Ginesu G., Pintus M., Giusto D.D. Objective assessment of the WebP image coding algorithm // Signal Processing: Image Communication. 2012. Vol. 27, no. 8. P. 867–874.
9. Исследование методов квантования латентного пространства вариационного автокодировщика для кадров FPV видеопотока. Часть II / А.А. Березкин [и др.] // Электросвязь. 2024. № 7. С. 16–25.
10. Jiang J., Zhang K., Timofte R. Towards flexible blind JPEG artifacts removal // IEEE/CVF International Conference on Computer Vision. 2021. P. 4997–5006.
11. Saveljev V., Kim S.K., Kim J. Moire effect in displays: A tutorial // Optical Engineering. 2018. Vol. 57, no. 3. URL: https://www.researchgate.net/publication/324074641_Moire_effect_in_displays_A_tutorial (дата обращения: 22.07.2024).
12. Zaitsev M., Maclaren J., Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions // Journal of Magnetic Resonance Imaging. 2015. Vol. 42, no. 4. P. 887–901.
13. Castellanos N.P., Makarov V.A. Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis // Journal of Neuroscience Methods. 2006. Vol. 158, no. 2. P. 300–312.
14. Wagenaar D.A., Potter S.M. Real-time multi-channel stimulus artifact suppression by local curve fitting // Journal of Neuroscience Methods. 2002. Vol. 120, no. 2. P. 113–120.
15. Deep generative adversarial compression artifact removal / L. Galteri [et al.] // IEEE International Conference on Computer Vision (ICCV). 2017. P. 4826–4835.
16. Cavigelli L., Hager P., Benini L. CAS-CNN: A deep convolutional neural network for image compression artifact suppression // 2017 International Joint Conference on Neural Networks (IJCNN). 2017. P. 752–759.

17. Deep learning-based compressed image artifacts reduction based on multi-scale image fusion / C.H. Yeh [et al.] // Information Fusion. 2021. Vol. 67, no. 4. P. 195–207.
18. Github. Stable diffusion: development repository. URL: <https://github.com/pesser/stable-diffusion/tree/main> (дата обращения: 28.07.2024).
19. Github. Lossy image compression with conditional diffusion models. URL: https://github.com/buggyyang/CDC_compression (дата обращения: 28.07.2024).

Получено 02.08.2024

Березкин Александр Александрович, к.т.н., доцент кафедры программной инженерии и вычислительной техники (ПИИВТ) Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ). 193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, 22-1. Тел. +7 921 791-61-20. E-mail: berezkin.aa@sut.ru

Ченский Александр Александрович, магистрант кафедры ПИИВТ СПбГУТ. 193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, 22-1. Тел. +7 911 091-35-57. E-mail: chenskii.aa@sut.ru

Киричек Руслан Валентинович, д.т.н., ректор, профессор кафедры ПИИВТ СПбГУТ. 193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, 22-1. Тел. +7 812 305-12-00. E-mail: kirichek@sut.ru

SUPPRESSION OF THE VIDEO STREAM FRAMES PROCESSED BY UNMANNED SYSTEMS USING FPV CONTROL

Berezkin A.A., Chenskiy A.A., Kirichek A.A.

*Bonch-Bruevich Saint Petersburg State University of Telecommunications,
Saint Petersburg, Russian Federation*

E-mail: berezkin.aa@sut.ru, chenskii.aa@sut.ru, kirichek@sut.ru

Using packet networks for first-person control of unmanned systems arises a problem of large transmitted data volumes. The largest volume of traffic during first-person control is presented by video stream frames. So, to improve the efficiency of the communication network between unmanned systems and external pilot station, it is necessary to compress video stream frames. A high compression degree can be provided by using variational autoencoders. One of the problems of using variational autoencoders for frame compression is the occurrence of specific artifacts in frames. This article proposes methods for suppressing the occurrence of artifacts when restoring frames from the latent space by a neural network decoder, as well as an empirical scale for assessing autoencoder artifacts. The approach proposed encompasses preparing pixel data of a video stream frame for encoding and further reconstruction after decoding. It is experimentally shown that one of the proposed methods allows eliminating the absolute majority of artifacts without introducing significant distortions into the reconstructed frames.

Keywords: *neural network, artifact suppression, frame artifacts, video stream transmission, variational autoencoder, neural codec, FPV-control, first person view control, unmanned systems, UAV, unmanned aircraft vehicles*

DOI: 10.18469/ikt.2024.22.1.01

Berezkin Alexander Alexandrovich, Bonch-Bruevich Saint Petersburg State University of Telecommunications, 22-1, Bolshevnikov Avenue, Saint Petersburg, 193232, Russian Federation; Associate Professor of Program Engineering and Computer Science Department, PhD in Technical Science. Tel. +7 921 791-61-20. E-mail: berezkin.aa@sut.ru

Chenskiy Alexander Alexandrovich, Bonch-Bruevich Saint-Petersburg State University of Telecommunications, 22-1, Bolshevnikov Avenue, Saint Petersburg, 193232, Russian Federation; Master's Degree Student of Program Engineering and Computer Science Department. Tel. +7 911 091-35-57. E-mail: chenskii.aa@sut.ru

Kirichek Ruslan Valentinovich, Bonch-Bruевич Saint Petersburg State University of Telecommunications, 22-1, Bolshhevikov Avenue, Saint Petersburg, 193232, Russian Federation; Rector, Professor of Program Engineering and Computer Science Department, Doctor of Technical Science. Tel. +7 812 305-12-00. E-mail: kirichek@sut.ru

References

1. Berezkin A.A. et al. Research of latent video stream compression methods for FPV control of uavs. *Elektrosvyaz*, 2024, no. 6, pp. 26–36. (In Russ.)
2. Berezkin A.A. et al. Research of latent space quantization methods of variational autoencoder for FPV video stream frames. Part I. *Elektrosvyaz*, 2024, no. 6, pp. 10–16. (In Russ.)
3. Project of the strategy for the development of the telecommunications industry of the Russian Federation for the period up to 2035 [adopted by the Government of Russian Federation on November 24, 2023]. URL: <https://digital.gov.ru/ru/documents/9120/> (accessed: 27.07.2024). (In Russ.)
4. ITU-T Recommendation H.264. Improved Image Encoding for General Audiovisual Services. Geneva, 2008, 342 p. (In Russ.)
5. ITU-T Recommendation H.265 (V9). High Efficiency Video Coding. Geneva, 2023, 718 p.
6. ISO/IEC 10918-1:1994. Information Technology – Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines. URL: <https://www.iso.org/ru/standard/18902.html> (accessed: 27.07.2024).
7. Barman N., Martini M.G. An evaluation of the next-generation image coding standard AVIF. *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020, pp. 1–4.
8. Ginesu G., Pintus M., Giusto D.D. Objective assessment of the WebP image coding algorithm. *Signal Processing: Image Communication*, 2012, vol. 27, no. 8, pp. 867–874.
9. Berezkin A.A. et al. Research of methods of quantization of latent space of variational autoencoder for FPV video stream frames. Part II. *Elektrosvyaz*, 2024, no. 7, pp. 16–25. (In Russ.)
10. Jiang J., Zhang K., Timofte R. Towards flexible blind JPEG artifacts removal. *IEEE/CVF International Conference on Computer Vision*, 2021, pp. 4997–5006.
11. Saveljev V., Kim S. K., Kim J. Moire effect in displays: A tutorial. *Optical Engineering*, 2018, vol. 57, no. 3. URL: https://www.researchgate.net/publication/324074641_Moire_effect_in_displays_A_tutorial (accessed: 22.07.2024).
12. Zaitsev M., Maclaren J., Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*, 2015, vol. 42, no. 4, pp. 887–901.
13. Castellanos N.P., Makarov V.A. Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods*, 2006, vol. 158, no. 2, pp. 300–312.
14. Wagenaar D.A., Potter S.M. Real-time multi-channel stimulus artifact suppression by local curve fitting. *Journal of Neuroscience Methods*, 2002, vol. 120, no. 2, pp. 113–120.
15. Galteri L. et al. Deep generative adversarial compression artifact removal. *IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4826–4835.
16. Cavigelli L., Hager P., Benini L. CAS-CNN: A deep convolutional neural network for image compression artifact suppression. *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 752–759.
17. Yeh C.H. et al. Deep learning-based compressed image artifacts reduction based on multi-scale image fusion. *Information Fusion*, 2021, vol. 67, no. 4, pp. 195–207.
18. Github. Stable diffusion: development repository. URL: <https://github.com/pesser/stable-diffusion/tree/main> (accessed: 28.07.2024).
19. Github. Lossy Image compression with conditional diffusion models. URL: https://github.com/buggyyang/CDC_compression (accessed: 28.07.2024).

Received 02.08.2024