

УДК 004.627, 004.932

DOI: 10.22213/2413-1172-2022-3-74-81

Метод оценки качества работы алгоритмов сжатия видео при передаче по низкоскоростному радиоканалу в условиях воздействия помех

А. А. Лянгузов, аспирант, ИжГТУ имени М. Т. Калашникова, Ижевск, Россия

А. В. Коробейников, кандидат технических наук, ИжГТУ имени М. Т. Калашникова, Ижевск, Россия

Представлен обзор работ, посвященных алгоритмам сжатия видео на основе методов искусственного интеллекта. Выделены два основных направления развития таких алгоритмов – создание модулей для постобработки результатов работы классических алгоритмов и разработка алгоритмов, полностью заменяющих существующие видеокodeки. Рассматриваются проблемы широко используемых критериев оценки качества работы алгоритмов сжатия и передачи видеоданных. Установлено, что критерии на базе вычисления среднеквадратичного отклонения, такие как, например, PSNR, не могут быть применены для оценки качества работы алгоритмов сжатия видеоданных в условиях воздействия радиотехнических помех ввиду сильного влияния на результат оценки артефактов, неизбежно возникающих в результате воздействия такого рода помех. Предложен альтернативный метод оценки качества на основе семантического анализа кадров, который может быть применен для оценки помехоустойчивости алгоритмов передачи видеоданных. Представленный метод оценки использует покaдровое сравнение исходного и восстановленного видеоряда аналогично PSNR. Однако в отличие от последнего игнорирует артефакты, возникающие в ходе воздействия радиотехнических помех, за счет использования семантического анализа кадров, заключающегося в поиске объектов на изображении с помощью подходов на основе искусственного интеллекта. Для сравнения найденных объектов на исходном и восстановленном кадре используется оценка, базирующаяся на геометрическом положении найденных объектов на изображении. После чего данный показатель усредняется для всех кадров видеоряда с целью получения результирующего показателя похожести восстановленного видео и оригинала.

Ключевые слова: сжатие видео, видеокodeки, обработка видео, оценка качества видео, помехоустойчивость передачи видео, семантический анализ видеоданных.

Введение

Сжатие данных – одна из технологий, благодаря которым стало возможным практическое использование потоковой передачи видеоданных. На сегодняшний день эта технология используется во многих аспектах современной жизни. С помощью потоковой передачи видеоданных работают прямые трансляции на телевидении и различных видеохостингах, осуществляются видеоконференции между людьми из разных стран или даже с разных континентов. Кроме того, данная технология используется для работы охранных систем, автопилотов, систем управления различной беспилотной техникой.

Одной из самых важных проблем, которую требуется решить для использования потоковой передачи видео, является искажение данных в процессе их передачи по каналу связи, вызванное различного рода помехами. На этапе декодирования при возникновении подобного рода ошибок зачастую крайне сложно восстановить исходный сигнал. Существующие на сегодняшний день видеокodeки обеспечивают хорошую степень сжатия видео и позволяют изба-

виться от существенного числа избыточной информации, создающей дополнительный трафик, а работа с помехами практически целиком возлагается на аппаратуру передачи по каналу связи. Современные средства коммуникации позволяют передавать данные на высокой скорости в широком диапазоне частот канала связи. За счет этого ошибки, возникающие под влиянием помех, зачастую не так критичны, так как могут быть устранены повторной отправкой сообщения. Однако существует ряд прикладных задач, где частоты канала связи строго регламентированы, а также присутствуют аппаратные ограничения на скорость передачи данных. В таких системах реализация потоковой передачи видео через повторную отправку кадров в условиях помех может давать существенную временную задержку, что делает данную технологию неприменимой в случае, когда время реакции на изменения во входных данных должно быть минимальным. В качестве примера можно привести задачу удаленного управления беспилотным летательным аппаратом.

Одним из возможных решений данной проблемы является использование методов искус-

ственного интеллекта. Машинное обучение позволяет выделить новые закономерности в данных при решении различных прикладных задач. За счет этого становится возможным еще большее сокращение объема трафика, необходимого для передачи видео или же использование найденных зависимостей на этапе декодирования сигнала для восстановления исходных видеоданных. Разработка помехоустойчивого алгоритма сжатия видео на основе машинного обучения позволила бы существенно сократить аппаратные требования для реализации потоковой передачи видео, существенно расширив тем самым область применения данной технологии.

Для решения озвученной в предыдущем разделе задачи требуется разработать алгоритм сжатия видео на основе методов искусственного интеллекта для его последующего применения при потоковой передаче видео по низкоскоростному каналу в условиях помех. Под низкоскоростным каналом в данной работе подразумевается радиоканал, аналогичный используемому в технологиях цифрового радиовещания Digital Radio Mondiale – цифровое глобальное радио (DRM), который обеспечивает скорость передачи данных от 37 кбит/с до 186 кбит/с (<https://www.drm.org/wp-content/uploads/2020/05/DRM-Handbook-Version-5.pdf>).

Цель данного исследования – разработка критерия оценки помехоустойчивости алгоритмов сжатия видео.

Объектом исследования являются алгоритмы сжатия данных, а предметом – помехоустойчивость этих алгоритмов.

Для достижения поставленной цели прежде всего требуется оценить применимость существующих решений. В данной работе приведен краткий обзор существующих методов сжатия видео на основе методов искусственного интеллекта, выделены основные направления исследований, а также представлен предлагаемый метод оценки качества работы алгоритмов сжатия видео с целью проверки их применимости для решения задачи потоковой передачи видео по низкоскоростному каналу в условиях помех.

Направления развития исследований

На сегодняшний день ведется множество исследований в области сжатия видеоданных с использованием технологий искусственного интеллекта [1–4]. Среди них можно выделить два основных направления развития.

Первое направление заключается в оптимизации работы существующих видеокодеков. Так, например, в работе [5] предлагается алгоритм, основанный на использовании сверточ-

ных нейронных сетей, работающий в совместно с видеокодеком H.264/AVC [6]. Основная идея предложенного алгоритма состоит в сокращении размера опорного кадра с его последующим восстановлением на стороне декодера. Представленные в работе результаты свидетельствуют об улучшении качества передаваемого изображения даже при низкой скорости передачи данных по сравнению с работой классической версии кодека H.264.

Похожий подход используется в статье [7], где авторы пытаются улучшить механизм межкадровой декорреляции, реализованный в стандарте HEVC [8], с использованием сверточной нейронной сети.

Еще один механизм, использующий межкадровую декорреляцию, представлен в статье [9]. В данной работе авторами был разработан многокадровый внутриконтурный фильтр «MIF» для стандарта HEVC. Данный фильтр улучшает качество каждого кадра, используя информацию о соседних с ним кадрах.

Другим направлением является полная замена существующих видеокодеков на их нейросетевые аналоги. Так, например, в работе [10] описывается алгоритм иерархического обучаемого сжатия видео (Hierarchical Learned Video Compression, HLVC) с тремя уровнями качества и рекуррентной сетью для его последующего улучшения. На этапе кодирования каждый кадр видео проходит через все три уровня сжатия. На первом слое кадры сжимаются с наилучшим сохранением качества, фактически методами сжатия без потерь. Далее ссылки на сжатые таким образом кадры используются для повторного сжатия с относительно высоким качеством на втором слое с помощью сети двунаправленного глубокого сжатия (BDDC). На третьем слое кадры сжимаются с самым низким качеством с помощью сети Single Motion Deep Compression (SMDC), которая использует единую карту для оценки движения нескольких кадров, тем самым экономя биты, кодирующие информацию о движении. В декодере, в свою очередь, используется сеть взвешенного рекуррентного улучшения качества (WRQE), с помощью которой алгоритм стремится восстановить утерянную при сжатии информацию.

Другой алгоритм, описанный в статье [11], предполагает полную замену кодирования кадров на генерацию кадров видео с использованием нейронных сетей. Предлагаемый метод кодирует ключевые кадры (I-кадры) как неподвижные изображения и полностью пропускает сжатие и передачу промежуточных кадров

(S-кадров), которые синтезируются исключительно с использованием I-кадров на стороне декодера за счет использования межкадровой взаимосвязи (корреляции). Результаты работы данного алгоритма позволяют предположить, что данная технология сжатия видео также является перспективной.

Однако у всех алгоритмов сжатия видео, основанных на методах искусственного интеллекта, есть некоторые общие проблемы. В работах [12, 13] поднимается проблема сложности прогнозирования движения с использованием нейронных сетей. Авторы предлагают концепцию Pixel-Motion CNN (PMCNN), способную моделировать пространственно-временную согласованность для эффективного выполнения прогнозирующего кодирования с помощью сверточной нейронной сети. Представленная сеть способна учитывать как межкадровые, так и внутрикадровые взаимосвязи (корреляции). Результаты экспериментов демонстрируют эффективность предложенной схемы и неплохие результаты работы в сравнении с кодеком H.264.

Кроме того, некоторые исследования современных авторов [14, 15] показали, что производительность работы нейросетевых видеокодеков зависит, в том числе, и от кодируемого содержимого. Иначе говоря, от объектов, присутствующих на конкретном видеоряде, будет зависеть эффективность сжатия. В указанной работе авторами представлена структура нейронной сети, позволяющей улучшить производительность нейросетевых кодеков. Предлагаемая сеть использует сходство между блоками изображения с помощью моделей гауссовой смеси без записи информации о сходстве с целью достижения лучшего качества реконструкции по сравнению с другими нейросетевыми алгоритмами.

Другой серьезной проблемой является вычислительная сложность работы нейросетевых моделей, используемых при сжатии видео. Чем больше параметров имеет модель, тем медленнее она работает. Данная проблема поднимается в работе [16], где авторы предлагают использовать особую предобработку данных, призванную сократить количество кадров при выполнении интерполяции исходного видеоряда.

Несмотря на это, исследователи компании WaveOne в своей работе [17] утверждают, что близки к революции в области сжатия видео. При обработке видео высокого разрешения 1080p их новый кодек, использующий алгоритмы машинного обучения, сжимает видео примерно на 20 % лучше, чем самые современные

традиционные видеокодеки, такие как H.265 и VP9. А на видео стандартной четкости (SD/VGA, 640×480) разница достигает 60 %. Основная идея сжатия видео заключается в удалении избыточных данных и замене их более коротким описанием. Для всего этого авторы применяют алгоритмы машинного обучения, которые по сравнению со стандартными алгоритмами позволяют найти новые зависимости между кадрами и, как следствие, избыточность в опорных кадрах. При всем этом получившийся алгоритм работает крайне медленно. На платформе Nvidia Tesla V100 при передаче видео стандартной четкости (SD) новый декодер работает со средней скоростью около 10 кадров в секунду, а кодер – со скоростью около 2 кадров в секунду. По этой причине данное решение крайне сложно применить в системах прямой видеотрансляции.

Критерий для оценки качества работы алгоритмов сжатия данных

Для оценки качества работы своих алгоритмов большинство разработчиков видеокодеков пользуется критерием «пиковое соотношение сигнал – шум» (PSNR) [18]. В общем случае критерий описывает соотношение между максимумами возможного значения сигнала и мощностью шума, искажающего значения сигнала, и является наиболее общепринятым критерием оценки различий между двумя последовательностями. Для изображения PSNR представляет собой логарифмическую оценку отношения максимального возможного значения, принимаемого пикселем изображения (при разрядности пикселей равной 8 битам максимально возможное значение будет равно 255), и среднеквадратичного отклонения значений пикселей исходного и восстановленного изображения.

Для двух изображений I и K размера $m \times n$ среднеквадратичное отклонение вычисляется по формуле

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2.$$

Тогда пиковое соотношение сигнала к шуму определяется как

$$PSNR = 20 \times \log_{10} \left(\frac{MAX_i}{\sqrt{MSE}} \right).$$

Однако данный критерий не лишен недостатков, основным из которых является способ оценки целевой величины. Критерий опирается

на точное соответствие результирующей последовательности и исходной. Благодаря этому при передаче видеосигнала существует ряд ситуаций, в которых критерий PSNR покажет существенное различие исходного и результирующего сигнала.

Пусть есть два кадра – исходный и результирующий (восстановленный после передачи). Один и тот же пиксель на исходном кадре белый, а на результирующем – черный. С точки зрения PSNR эти кадры существенно различаются, так как отсутствует точное соответствие исходной и результирующей последовательности.

Аналогичная ситуация наблюдается при изменении яркости результирующего изображения или видеокadra по сравнению с исходным. В этом случае PSNR будет так же сигнализировать о различии последовательностей. Данный недостаток не позволяет пользоваться этим критерием при оценке применимости алгоритма

сжатия данных для решения задачи потоковой передачи видео по низкочастотному каналу в условиях помех ввиду того, что точного соответствия последовательностей в данном случае не требуется. При решении этой задачи на результирующем кадре необходимо обнаруживать те же объекты, которые были на исходном кадре и на тех же позициях в кадре. Зашумленность результирующего видеоряда допустима до тех пор, пока объекты остаются визуально различимыми. Геометрические размеры объектов и их позиции также не должны искажаться.

Таким образом, для оценки качества работы алгоритмов сжатия данных при решении данной задачи предлагается ввести новый критерий и использовать его в дальнейших исследованиях. Его вычисление так же, как в PSNR, основано на сравнении исходного и результирующего видеоряда. Схема работы предлагаемого метода оценки приведена на рисунке 1.

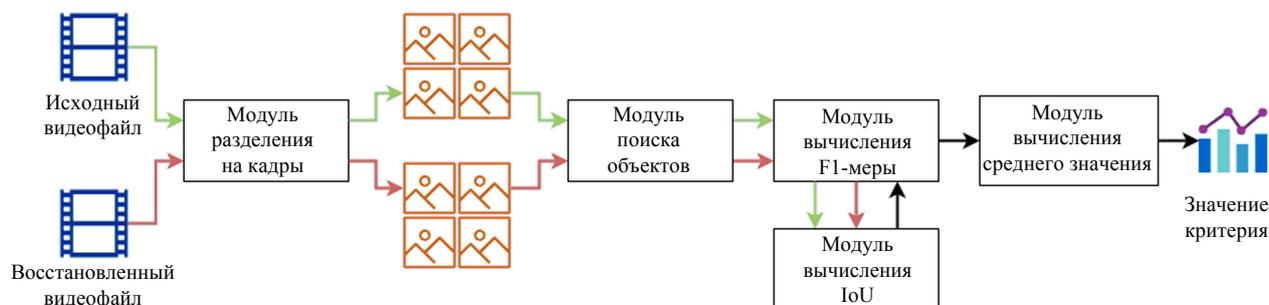


Рис. 1. Схема работы предлагаемого критерия оценки

Fig. 1. Proposed evaluation criterion scheme

На вход алгоритму оценки качества сжатия видео подается исходный и восстановленный после передачи видеофайлы.

Первым этапом работы алгоритма является разделение исходного и восстановленного видеофайлов на кадры. После этого выполняется семантический анализ каждого кадра исходного и восстановленного видеофайла, заключающийся в поиске объектов на каждом кадре. Для решения задачи поиска объектов на изображении предлагается использовать алгоритмы FasterR-CNN (<https://github.com/facebookresearch/Detectron>) или YOLO (<https://github.com/ultralytics/yolov3>), так как эти алгоритмы хорошо зарекомендовали себя и на сегодняшний день являются эталонами (State of the Art) при решении подобных задач. Пример поиска объектов алгоритмом FasterR-CNN представлен на рисунке 2.

Скорость распознавания непосредственно для оценки качества работы значения не имеет, поэтому на текущий момент основным алгоритмом семантического анализа определим

FasterR-CNN, а YOLO определим как дополнительный алгоритм.

Далее для каждого кадра исходного видео выполняется поиск соответствующего кадра на восстановленном видео, после чего формируется пара коллекций объектов, найденных на исходном и на восстановленном кадре.

После этого для каждой пары кадров выполняется сравнение параметров обнаруженных объектов на исходном и восстановленном кадре. Сравнение происходит по критерию F1-мера [19], вычисляемому как среднее гармоническое между критериями точности (precision) и полноты (recall), которые вычисляются с использованием матрицы ошибок, приведенной в таблице, и позволяют оценить ошибки первого и второго рода при оценке качества распознавания объектов алгоритмом. Точность (precision) позволяет оценить отношение верно распознанных объектов к общему числу распознанных объектов, а полнота (recall), в свою очередь, позволяет оценить долю верно распознанных объектов по

отношению ко всем объектам, которые необходимо было распознать.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall};$$

$$Recall = \frac{TP}{TP + FN};$$

$$Precision = \frac{TP}{TP + FP};$$

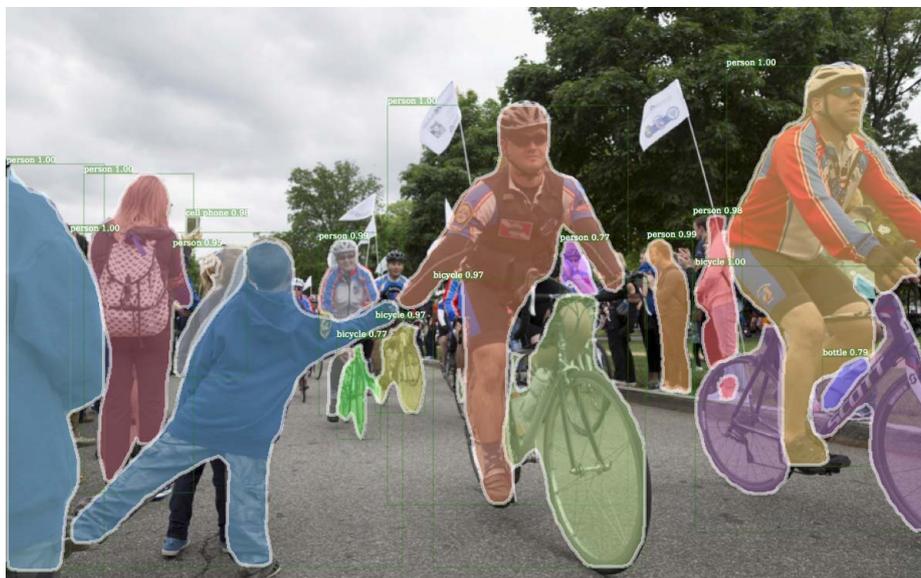


Рис. 2. Пример работы алгоритма FasterR-CNN

Fig. 2. An example of the FasterR-CNN algorithm result

Матрица ошибок

Error Matrix

Предсказанные значения / Истинные значения	Y = 1	Y = 0
$\bar{Y} = 1$	True Positive (TP)	False Positive (FP)
$\bar{Y} = 0$	FalseNegative (FN)	TrueNegative (TN)

Для оценки геометрического совпадения двух объектов предлагается использовать критерий Intersection overUnion (IoU), приведенный в работе [20] и представляющий собой отношение площади пересечения объектов к площади их объединения, с пороговым значением 0,8.

Формула для расчета представлена на рисунке 3.

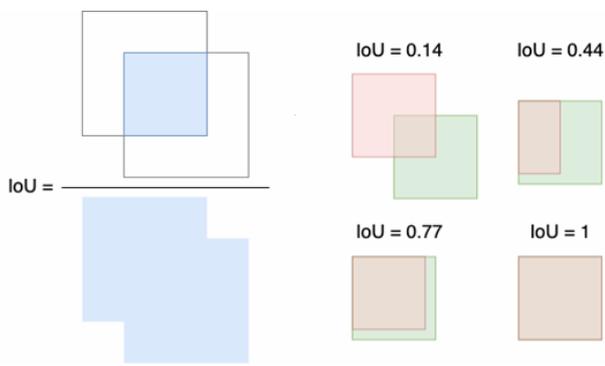


Рис. 3. Формула для расчета IoU

Fig. 3. IoU calculation formula

Использование такого критерия позволяет корректно учитывать границы объектов, их положение и размеры, не привязываясь к абсолютным координатам. При необходимости пороговое значение может быть скорректировано. На текущий момент оно подобрано интуитивно, исходя из опыта решения задач по извлечению именованных сущностей из текстовых данных (<https://www.directum.ru/products/arlio>), где через IoU часто сравниваются положения текстовых сегментов на странице документа.

На следующем этапе работы алгоритма для каждой пары коллекций найденных объектов фиксируется значение F1-меры. Фактически происходит семантический анализ кадра, учитывающий ошибки распознавания, размеры и позиции объектов в кадре.

Финальным этапом является вычисление среднего значения F1-меры по всем парам кадров исходного и восстановленного видеорядов. В качестве среднего значения предлагается брать среднее гармоническое по аналогии

с классической формулой для расчета F1-меры. Именно эту величину предлагается использовать в качестве критерия, определяющего качество работы алгоритма сжатия данных. Данный критерий имеет ту же область значений, что и F1-мера: $[0,0; 1,0]$, так как в действительности является ее вариацией. При полном соответствии исходной и результирующей последовательности значение критерия будет 1,0.

Разработанный критерий можно использовать не только для оценки качества работы алгоритма в идеальных условиях, но также для измерения его помехоустойчивости. В том числе с помощью него можно оценить устойчивость алгоритма к кратковременной потере связи, что также является важной характеристикой при решении задачи потоковой передачи видео.

Следует отметить, что предложенный критерий требует полного набора пар кадров, исходного и восстановленного видеорядов. В случае отсутствия в восстановленном видеоряде кадров можно считать восстановленный кадр пустым (не содержащим объектов). В этом случае при сравнении исходного кадра с пустым восстановленным кадром результат оценки по критерию F1-мера будет 0,1.

Заключение

Разработан метод оценки качества работы алгоритмов сжатия данных на основе семантического анализа кадров, который позволяет не только оценить идентичность исходных и восстановленных видеоданных, но быть уверенным в том, что видео остается пригодным для распознавания объектов человеком или программной системой, невзирая на помехи, возникшие при его передаче.

Выводы

1) Современные широко используемые методы оценки качества работы алгоритмов сжатия видео не могут быть применены для оценки помехоустойчивости этих алгоритмов, так как наличие артефактов, возникающих на этапе передачи, под влиянием помех приводит к существенному различию исходных и восстановленных видеоданных, хотя видео по-прежнему может оставаться пригодным для дальнейшей машинной обработки или восприятия человеком.

2) Оценка помехоустойчивости алгоритмов сжатия данных может быть проведена с использованием семантического анализа кадров.

Предложенные в статье положения имеют значение для использования на практике в сфере передачи видеоданных в условиях наличия радиотехнических помех с целью оценки помехоустойчивости алгоритмов сжатия видео.

Применение разработанного метода оценки позволит получить оценку качества восстановленного видео, которая может быть использована для оценки эффективности передачи данных в условиях помех.

Библиографические ссылки

1. *Ma S.* Image and video compression with neural networks: A review: IEEE Transactions on Circuits and Systems for Video Technology, 2019, vol. 30, no. 6, pp. 1683-1698.
2. *Liu D.* Deep learning-based technology in responses to the joint call for proposals on video compression with capability beyond HEVC: IEEE Transactions on Circuits and Systems for Video Technology, 2019, vol. 30, no. 5, pp. 1267-1280.
3. *Cheng Z., Sun H., Takeuchi M., Katto J.* Learned image compression with discretized gaussian mixture likelihoods and attention modules: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
4. *Ding D.* Advances in video compression system using deep neural network: A review and case studies: Proc. of the IEEE, 2021, vol. 109, no. 9, pp. 1494-1520.
5. *Gorodilov A., Gavrilov D., Schelkunov D.* Neural Networks for Image and Video Compression. International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI), 2018, pp. 37-41. URL: <https://ieeexplore.ieee.org/document/8674448> (дата обращения: 10.01.2022).
6. *Wiegand T.* Overview of the H. 264/AVC video coding standard: IEEE Transactions on circuits and systems for video technology, 2003, vol. 13, no. 7, pp. 560-576.
7. *Li Y.* Convolutional Neural Network-Based Block Up-Sampling for Intra Frame Coding: IEEE Transactions on Circuits and Systems for Video Technology, 2018, vol. 28, no. 9, pp. 2316-2330. DOI: 10.1109/TCSVT.2017.2727682.
8. *Sullivan G.J.* Overview of the high efficiency video coding (HEVC) standard: IEEE Transactions on circuits and systems for video technology, 2012, vol. 22, no. 12, pp. 1649-1668.
9. *Li T.* A deep learning approach for multi-frame in-loop filter of HEVC: IEEE Transactions on Image Processing, 2019, vol. 28, no. 11, pp. 5663-5678.
10. *Ren Yang, Fabian Mentzer, Luc Van Gool, Radu Timofte.* Learning for Video Compression With Hierarchical Quality and Recurrent Enhancement: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 6628-6637. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Yang_Learning_for_Video_Compression_With_Hierarchical_Quality_and_Recurrent_Enhancement_CVPR_2020_paper.html (дата обращения: 10.01.2022).
11. *Nicola Giuliani, Biao Wang, Elena Alshina, Laura Leal-Taixé.* Frame synthesis for video compression: Proc. SPIE 11842, Applications of Digital Image Proc. XLIV, 118420R (1 August 2021). URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11842/2596884/Frame-synthesis->

for-video-compression/10.1117/12.2596884.short?SSO=1 (дата обращения: 10.01.2022).

12. Chen Z., He T., Jin X., Wu F. Learning for Video Compression: IEEE Transactions on Circuits and Systems for Video Technology, 2020, vol. 30, no. 2, pp. 566-576. URL: <https://ieeexplore.ieee.org/abstract/document/8610323> (дата обращения: 10.01.2022).

13. Lee H., Kim T., Chung T-young., Pak D., Ban Y., Lee S. Adaptive collaboration of flows for video frame interpolation: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5316–5325.

14. Pei Y., Liu Y., Ling N., Liu L., Ren Y. Class-Specific Neural Network for Video Compressed Sensing: IEEE International Symposium on Circuits and Systems (ISCAS), 2021, pp. 1-5. URL: <https://ieeexplore.ieee.org/document/9401450> (дата обращения: 10.01.2022).

15. Bao W., Lai W.-S., Ma C., Zhang X., Gao Z., Yang M.-H. Depth-aware video frame interpolation: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 3703-3712.

16. Choi M. Motion-aware dynamic architecture for efficient frame interpolation: Proc. of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 13839-13848.

17. Oren Rippel, Sanjay Nair, Carissa Lew, Steve Branson, Alexander G. Anderson, Lubomir Bourdev. Learned Video Compression. URL: <https://doi.org/10.48550/arXiv.1811.06981> (дата обращения: 21.02.2022).

18. Welstead Stephen T. Fractal and wavelet image compression techniques. *SPIE Publication, 2019*, pp. 155-156. ISBN 978-0-8194-3503-3 (дата обращения: 02.06.2022).

19. Erik F. Tjong Kim Sang and Fien De Meulder. Introduction to the CoNLL-2003 shared task: language-independent named entity recognition: Proc. of the seventh conference on Natural language learning at HLT-NAACL, 2003, vol. 4 (CONLL '03). Association for Computational Linguistics, USA, pp. 142-147. <https://doi.org/10.3115/1119176.1119195> (дата обращения 02.06.2022).

20. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation: CVPR, 2014. URL: <https://doi.org/10.48550/arXiv.1311.2524>.

References

Ma S. Image and video compression with neural networks: A review: IEEE Transactions on Circuits and Systems for Video Technology, 2019, vol. 30, no. 6, pp. 1683-1698.

2. Liu D. Deep learning-based technology in responses to the joint call for proposals on video compression with capability beyond HEVC: IEEE Transactions on Circuits and Systems for Video Technology, 2019, vol. 30, no. 5, pp. 1267-1280.

3. Cheng Z., Sun H., Takeuchi M., Katto J. Learned image compression with discretized gaussian mixture likelihoods and attention modules: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

4. Ding D. Advances in video compression system using deep neural network: A review and case studies: Proc. of the IEEE, 2021, vol. 109, no. 9, pp. 1494-1520.

5. Gorodilov A., Gavrilo D., Schelkunov D. Neural Networks for Image and Video Compression. International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI), 2018, pp. 37-41. URL: <https://ieeexplore.ieee.org/document/8674448> (дата обращения: 10.01.2022).

6. Wiegand T. Overview of the H. 264/AVC video coding standard: IEEE Transactions on circuits and systems for video technology, 2003, vol. 13, no. 7, pp. 560-576.

7. Li Y. Convolutional Neural Network-Based Block Up-Sampling for Intra Frame Coding: IEEE Transactions on Circuits and Systems for Video Technology, 2018, vol. 28, no. 9, pp. 2316-2330. DOI: 10.1109/TCSVT.2017.2727682.

8. Sullivan G.J. Overview of the high efficiency video coding (HEVC) standard: IEEE Transactions on circuits and systems for video technology, 2012, vol. 22, no. 12, pp. 1649-1668.

9. Li T. A deep learning approach for multi-frame in-loop filter of HEVC: IEEE Transactions on Image Processing, 2019, vol. 28, no. 11, pp. 5663-5678.

10. Ren Yang, Fabian Mentzer, Luc Van Gool, Radu Timofte. Learning for Video Compression With Hierarchical Quality and Recurrent Enhancement: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 6628-6637. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Yang_Learning_for_Video_Compression_With_Hierarchical_Quality_and_Recurrent_Enhancement_CVPR_2020_paper.html (дата обращения: 10.01.2022).

11. Nicola Giuliani, Biao Wang, Elena Alshina, Laura Leal-Taixé. Frame synthesis for video compression: Proc. SPIE 11842, Applications of Digital Image Proc. XLIV, 118420R (1 August 2021). URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11842/2596884/Frame-synthesis-for-video-compression/10.1117/12.2596884.short?SSO=1> (дата обращения: 10.01.2022).

12. Chen Z., He T., Jin X., Wu F. Learning for Video Compression: IEEE Transactions on Circuits and Systems for Video Technology, 2020, vol. 30, no. 2, pp. 566-576. URL: <https://ieeexplore.ieee.org/abstract/document/8610323> (дата обращения: 10.01.2022).

13. Lee H., Kim T., Chung T-young., Pak D., Ban Y., Lee S. Adaptive collaboration of flows for video frame interpolation: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5316–5325.

14. Pei Y., Liu Y., Ling N., Liu L., Ren Y. Class-Specific Neural Network for Video Compressed Sensing: IEEE International Symposium on Circuits and Systems (ISCAS), 2021, pp. 1-5. URL: <https://ieeexplore.ieee.org/document/9401450> (дата обращения: 10.01.2022).

15. Bao W., Lai W.-S., Ma C., Zhang X., Gao Z., Yang M.-H. Depth-aware video frame interpolation: Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 3703-3712.

16. Choi M. Motion-aware dynamic architecture for efficient frame interpolation: Proc. of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 13839-13848.

17. Oren Rippel, Sanjay Nair, Carissa Lew, Steve Branson, Alexander G. Anderson, Lubomir Bourdev. Learned Video Compression. URL: <https://doi.org/10.48550/arXiv.1811.06981> (дата обращения: 21.02.2022).

18. *Welstead Stephen T.* Fractal and wavelet image compression techniques. *SPIE Publication, 2019*, pp. 155-156. ISBN 978-0-8194-3503-3 (дата обращения: 02.06.2022).

19. Erik F. Tjong Kim Sang and Fien De Meulder. Introduction to the CoNLL-2003 shared task: language-independent named entity recognition: Proc. of the seventh conference on Natural language learning at HLT-NAACL, 2003, vol. 4 (CONLL '03). Association for Computational Linguistics, USA, pp. 142-147. <https://doi.org/10.3115/1119176.1119195> (дата обращения 02.06.2022).

20. Girshick R., Donahue J., Darrell T. Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation: CVPR, 2014. URL: <https://doi.org/10.48550/arXiv.1311.2524>.

Video Compression Performance Evaluation Method in Transmission via a Low-Speed Channel

A.A. Lyanguzov, Post-graduate, Kalashnikov ISTU, Izhevsk, Russia

A.V. Korobeynikov, PhD in Engineering, Associate Professor, Kalashnikov ISTU, Izhevsk, Russia

The paper presents an overview of the works devoted to video compression algorithms based on artificial intelligence methods. Two main directions for the development of such algorithms are identified - these are the development of modules for post-processing the results of the work of classical algorithms and the development of algorithms that completely replace existing video codecs. The problems of video compression algorithms performance evaluation criteria are considered. It was found that criteria based on the calculation of the standard deviation, such as PSNR, cannot be used to assess the quality of video data compression algorithms when operating under the influence of radio interference, due to the strong influence on the evaluation result of artifacts that inevitably arise as a result of the effects of this kind of interference. An alternative quality assessment method based on the semantic analysis of frames is proposed, which can be used to assess the noise immunity of video data transmission algorithms. The presented estimation method uses a frame-by-frame comparison of the original and restored video sequence, similar to PSNR. However, unlike the latter, it ignores artifacts that occur during exposure to radio interference, due to the use of semantic frame analysis, which consists in searching for objects in the image using approaches based on artificial intelligence. To compare the found objects on the original and reconstructed frames, an estimate is used based on the geometric position of the found objects in the image. After that, the data indicator is averaged for all frames of the video sequence in order to obtain the resulting similarity indicator of the restored video and the original.

Keywords: video compression, video codecs, video processing, video compression quality evaluation, interference resistance of video compression, video frames semantic analysis.

Получено 14.06.2022

Образец цитирования

Лянгузов А. А., Коробейников А. В. Метод оценки качества работы алгоритмов сжатия видео при передаче по низкоскоростному радиоканалу в условиях воздействия помех // Вестник ИжГТУ имени М. Т. Калашникова. 2022. Т. 25, № 3. С. 74–81. DOI: 10.22213/2413-1172-2022-3-74-81.

For Citation

Lyanguzov A.A., Korobeynikov A.V. [Video Compression Performance Evaluation Method in Transmission via a Low-Speed Channel]. *Vestnik IzhGTU imeni M.T. Kalashnikova*, 2022, vol. 25, no. 3, pp. 74-81 (in Russ.). DOI: 10.22213/2413-1172-2022-3-74-81.