

что определяет условие как обучаемые контекстуальные функции.

Для генерации полезного контекста совершается переход из пространственной области в пространство признаков. Для построения векторов движения между входным и предыдущим декодированным кадром используется нейросеть оценки оптического потока. Полученная после кодирования и декодирования вектора движения информация определяет, где извлечь контекст. В то же время с помощью нейросети осуществляется выделение признаков из предыдущего декодированного кадра. Применяя к полученной информации операцию деформации получим относительно грубый контекст, после чего производится уточнение контекста, реализованное на основе машинного обучения.

С использованием уточненного контекста в контекстном кодере входной кадр кодируется и затем квантуется через операцию округления. Для избавления от избыточности на этом этапе применяются алгоритмы машинного обучения. Также уточненный контекст используется в реализации энтропийной модели. После декодирования с применением уточненного контекста на выходе получается декодированный кадр.

Б. Энтропийная модель

Энтропийная модель «Рис. 2» включает в себя три приора (иерархический, пространственный и временной).



Рис. 2. Схема энтропийной модели, используемой в методе глубокого контекстуального сжатия видео

Для учета временной корреляции обучается временной приор с применением определенного несколькими шагами ранее контекста. Все используемые

приоры объединяются сетью слияния, которая также позволяет оценить среднее и масштаб распределения скрытого представления. Эти параметры передаются в арифметический кодер и декодер для формирования массива сжатых данных.

3. ЭКСПЕРИМЕНТ

Проведены вычислительные эксперименты для исследования эффективности рассмотренного нейросетевого метода контекстуального сжатия видео на наборе видеопоследовательностей [4].

Зависимость показателя эффективности сжатия (PSNR) от степени сжатия, выраженной в бит на пиксель (bpp), приведена на «Рис. 3». Приведены также результаты эффективности работы стандартных видеокодеков H.264, H.265 и VP9 в качестве базы для сравнения. Вычислительные эксперименты продемонстрировали выигрыш рассматриваемого алгоритма глубокого контекстуального сжатия видео.

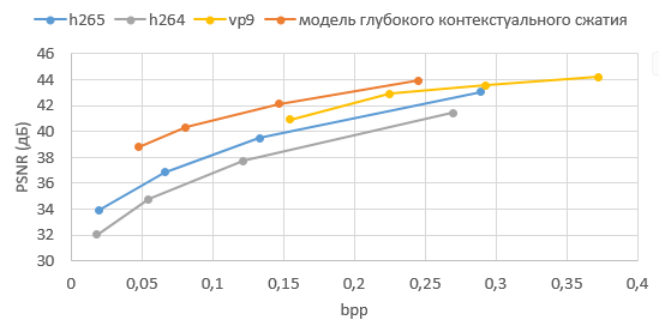


Рис. 3. Исследование эффективности модели глубокого контекстуального сжатия видео на основе машинного обучения

ЗАКЛЮЧЕНИЕ

Исследовался метод глубокого контекстуального сжатия видео на основе машинного обучения. Были приведены результаты вычислительных экспериментов на реальных тестовых видеопоследовательностях. Вычислительные эксперименты продемонстрировали выигрыш рассматриваемого метода на реальных видеопоследовательностях в сравнении с эффективностью стандартных видеокодеков.

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда (проект № 22-21-00662).

ЛИТЕРАТУРА

- [1] Ding, D. Advances in Video Compression System Using Deep Neural Network: A Review and Case Studies / D. Ding, Z. Ma, D. Chen, Q. Chen, Z. Liu, F. Zhu // Proceedings of the IEEE. – 2021. – Vol. 9(109). – P. 1494–1520.
- [2] Rippel, O. Learned video compression / O. Rippel, S. Nair, C. Lew, S. Branson, A. G. Anderson, L. Bourdev // Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2019. – P. 3454–3463.
- [3] Li, J. Deep Contextual Video Compression / J. Li, B. Li, Y. Lu // Advances in Neural Information Processing Systems. – 2021. – Vol. 34. – P. 18114–18125.
- [4] Набор видеороликов на основе пользовательского контента YouTube [Электронный ресурс]. – Режим доступа: <https://media.withyoutube.com/> (01.11.2022).