

# АЛГОРИТМ НАХОЖДЕНИЯ ПАРЕТО-ОПТИМАЛЬНЫХ КОНФИГУРАЦИЙ ВИДЕОКОДЕКА H264 ДЛЯ ПРОИЗВОЛЬНОЙ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

**Казанцев Роман Александрович,  
Звездаков Сергей Васильевич**

*Студент, Аспирант*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: roman.kazantsev@graphics.cs.msu.ru,*

*sergey.zvezdakov@graphics.cs.msu.ru*

**Научный руководитель — Ватолин Дмитрий Сергеевич**

Видеокодек h264 имеет 49 опций для настройки кодирования видео. Каждая конфигурация (настройка опций) видеокодека будет кодировать видеопоследовательность с разной скоростью и степенью искажения качества. Поэтому возникает задача многокритериальной оптимизации конфигурации видеокодека в пространстве степени искажения качества и скорости кодирования для заданного видео:

$$\operatorname{argmin}_{p \in S} (D_V(p), T_V(p)), \quad (1)$$

где  $S$  — множество всех (или рассматриваемых) настроек видеокодека h264,  $D_V(p)$  и  $T_V(p)$  — функции средних значений искажения качества и временных затрат конфигурации  $p$  для видео  $V$ . Решение задачи 1 для видеопоследовательности  $V$  обозначим через  $P_0(V)$ , которое будет состоять из множества Парето-оптимальных конфигураций.

Перед кратким описанием процедуры решения задачи 1 введем понятие структуры Парето-фронтон для видеопоследовательности. Из оставшихся конфигураций из множества  $S$  сформируем Парето-фронтонты  $P_i(V) = \operatorname{argmin}_{p \in S \setminus \cup_{k=0}^{i-1} P_k(V)} (D_V(p), T_V(p))$  с номером  $i$ , где  $i = 1, \dots, l$  и  $l$  — номер самого дальнего Парето-фронтонты.  $P_i(V)$ , где  $i = 0, \dots, l$ , будут использоваться для задания структуры Парето-фронтонты видеопоследовательности  $V$  в виде  $P_S(V) = (x_1, x_2, \dots, x_{|S|})$ , где  $x_k$  — номер Парето-фронтонты, которому принадлежит  $k$ -ая конфигурация из упорядоченного множества  $S$ .

Задача 1 была решена с помощью методов машинного обучения. Для этого были отобраны 351 видеопоследовательности и 1306 конфигурации видеокодека. Для каждой пары конфигурации и видеопоследовательности были вычислены средние значения искажения качества видео и скорости кодирования. Вычисления производились

около 3 недель на 145 компьютерах. По полученным данным были вычислены структуры Парето-фронт  $P_S(V)$  для каждой видеопоследовательности  $V$ . Видеопоследовательности были кластеризованы на классы по схожести их структур. В качестве меры схожести структур видеопоследовательностей  $V_1$  и  $V_2$  была использована корреляция Спирмена. Таким образом корреляция структур  $P_S(V_1) = (x_1, x_2, \dots, x_{|S|})$  и  $P_S(V_2) = (y_1, y_2, \dots, y_{|S|})$  будет вычисляться по формуле:  $\rho(V_1, V_2) = 1 - \frac{6}{|S|(|S|-1)(|S|+1)} \sum_{i=1}^{|S|} (x_i - y_i)^2$ . Видеопоследовательности были отнесены к одному классу, если  $\rho(V_1, V_2) \geq 0.9$ . Таким образом было получено 7 классов структур и задача 1 была сведена к задаче классификации.

Для того чтобы прогнозировать класс структур Парето-фронт заданной видеопоследовательности, из работ [1–3] были отобраны физические признаки видео, описывающие пространственную и временную сложность видео, четкость, среднее значение и разброс компонент пикселя, распределение цвета в кадре. Также мы сконструировали признак для оценки характера движения объектов. Признак представляет собой двумерную гистограмму движения в кадре, оценивающую распределение длины и направления (угла) вектора движения.

Для прогнозирования класса структур Парето-фронт по физическим признакам видео была обучена логистическая регрессия. Точность полученного классификатора оценивалась с помощью кросс-валидации и составила 75.2%.

В заключение авторы благодарны факультету ВМК МГУ за предоставленные вычислительные мощности для сбора соответствующих данных по скорости и качеству кодирования видеокodeка x264.

### Литература

1. Хабибуллина Н. А. Разработка новых методов анализа качества видеокodeков и оптимизация систем сжатия видеоинформации. Диссертация на соискание ученой степени кандидата технических наук. МФТИ, 2014.
2. Crete F. The Blur Effect: Perception and Estimation with a New No-Reference Perceptual Blur Metric // Human Vision and Electronic Imaging XII, 2007, P. 1–11.
3. Zumer R. Color-independent classification of animation video // International Journal of Multimedia Information Retrieval, 2018, P. 187–196.