

# Ортогонализация и параметризация свёрточных ядер в машинном обучении при компрессии изображений и видео

Р.Р. Юзькив

Самарский национальный исследовательский университет  
им. академика С.П. Королева  
Самара, Россия  
yuzkiv@geosamara.ru

М.В. Гашников

Самарский национальный исследовательский университет  
им. академика С.П. Королева  
Самара, Россия  
mih-fastt@yandex.ru

**Аннотация**—Исследуется ортогонализация и параметризация свёрточных ядер, применяемых при иерархической интерполяции в рамках метода компрессии изображений и видео, основанного на машинном обучении. В качестве алгоритмов машинного обучения на этапе интерполяции рассматриваются нейросети сверхразрешения и деревья решений, адаптивно выбирающие интерполирующую функцию из заранее заданного набора свёрточных функций с параметризованными ортогональными весами. Ортогонализация и параметризация весов свёрточных ядер используется для повышения эффективности интерполяционного алгоритма машинного обучения, которое в свою очередь приводит к повышению эффективности метода компрессии в целом. Вычислительные эксперименты демонстрируют выигрыш предложенного алгоритма на реальных видеосигналах.

**Ключевые слова**—ортогонализация, параметризация, видео, компрессия, машинное обучение.

## 1. ВВЕДЕНИЕ

Актуальность исследований в области использования методов машинного обучения для сжатия цифровых изображений и видеопоследовательностей [1, 2] опирается на высокую эффективность этих методов, позволяющих добиться кратного улучшения показателей качества во многих прикладных задачах.

Рассматриваемый метод компрессии развивает идею пирамидальной компрессии, которая основывается на мультимасштабном представлении (см. рис. 1) цифрового изображения в виде набора в разной степени прореженных уровней, позволяющем интерполировать пиксели на основе более прореженных уровней и кодировать постинтерполяционные остатки.

В этом методе компрессии алгоритмы машинного обучения используются, в частности, на этапе интерполяции. Наиболее подходящими методами машинного обучения на этапе интерполяции являются нейросети сверхразрешения и деревья решений, адаптивно выбирающие интерполирующую функцию из заранее заданного набора «атомарных» функций в каждой точке изображения или видео.

В качестве таких атомарных интерполирующих функций в данной работе рассматриваются свёртки. Ортогонализация и параметризация весов свёрточных ядер используется для повышения эффективности интерполяционного алгоритма машинного обучения, которое в свою очередь приводит к повышению эффективности метода компрессии в целом.

## 2. ИНТЕРПОЛЯЦИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ ПРИ СЖАТИИ ИЗОБРАЖЕНИЙ И ВИДЕО

В [3] автором настоящей работы предложен основанный на машинном обучении метод компрессии изображений, использующий современные нейросетевые алгоритмы сверхразрешения и доопределения цифровых изображений. Общая структура этого метода компрессии изображений основана на использовании алгоритмов генерации изображений, пирамидальной компрессии, и иерархической интерполяции (см. рис. 2).

Результаты экспериментов [3] позволили сделать вывод как о перспективности собственно метода компрессии, так и об эффективности алгоритмов машинного обучения, выбранных для использования в рамках этого метода.

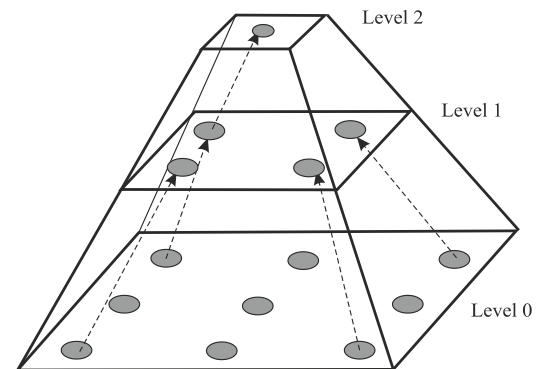


Рис. 1. Мультимасштабное представление изображения в виде набора из трёх прореженных уровней (с метками Level2, Level1, Level0)

## 3. ОРТОГОНАЛИЗАЦИЯ И ПАРАМЕТРИЗАЦИЯ СВЁРТОЧНЫХ ЯДЕР

Ортогонализация и параметризация свёрточных ядер в данной работе используется на этапе иерархической интерполяции (см. рис. 2). Одним из самых простых свёрточных ядер при интерполяции изображения или видеокadra по прореженному вдвое изображению или видеокadру является следующее ядро:

$$\begin{bmatrix} 1/4 & 0 & 1/4 \\ 0 & 0 & 0 \\ 1/4 & 0 & 1/4 \end{bmatrix} \quad (1)$$

Систему из двух параметризованных ортогональных свёрточных ядер для той же ситуации можно записать в виде:

$$\begin{bmatrix} a/2 & 0 & (1-a)/2 \\ 0 & 0 & 0 \\ (1-a)/2 & 0 & a/2 \end{bmatrix}, \quad (2)$$

$$\begin{bmatrix} (a-1)/(4a-2) & 0 & a/(4a-2) \\ 0 & 0 & 0 \\ a/(4a-2) & 0 & (a-1)/(4a-2) \end{bmatrix}, \quad (3)$$

где  $a$  – параметр из интервала  $[0, 0,5)$ .

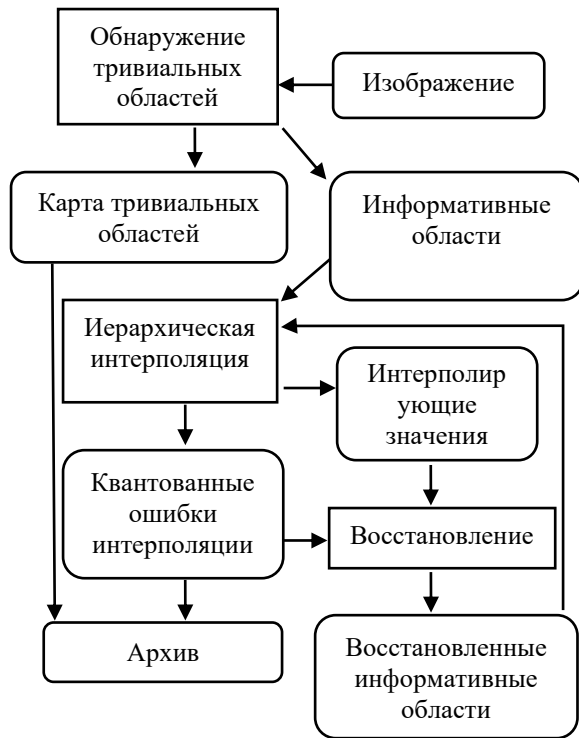


Рис. 2. Место иерархической интерполяции, использующей ортогональные параметризованные свёрточные ядра, в структуре основанного на машинном обучении метода компрессии изображений и видео

Для переключения между этими свёрточными ядрами можно использовать алгоритм машинного обучения на основе дерева решений. Обучение дерева решений можно выполнять непосредственно на сжимаемом изображении или видео, а вычисленные параметры записать в архивный файл, так как они понадобятся при декомпрессии.

#### 4. ЭКСПЕРИМЕНТ

Для исследования эффективности описанных ортогональных параметризованных свёрточных ядер при компрессии были проведены вычислительные эксперименты на видеонаборе [4] (см. пример тестового видеосигнала на рис. 3).

Показателем эффективности являлся относительный выигрыш в объёме сжатых данных, возникающий при замене ядра (1) на систему ортогональных параметризованных ядер (2-3). Усреднённая зависимость этого выигрыша от вносимой в видеоданные абсолютной

погрешности приведена на рис. 4. Вычислительные эксперименты продемонстрировали выигрыш предложенного алгоритма при практически значимых значениях абсолютной погрешности.



Рис. 3. Пример фрагмента тестовой видеопоследовательности

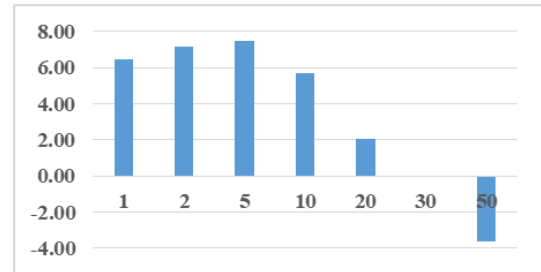


Рис. 4. Процент выигрыша в размере архива за счёт использования межкадровых взаимозависимостей в зависимости от максимальной погрешности

#### ЗАКЛЮЧЕНИЕ

Исследовалась ортогонализация и параметризация свёрточных ядер, применяемых при иерархической интерполяции в рамках метода компрессии изображений и видео, основанного на машинном обучении. В качестве алгоритмов машинного обучения на этапе интерполяции рассматривались нейросети сверхразрешения и деревья решений, адаптивно выбирающие интерполирующую функцию из заранее заданного набора свёрточных функций с параметризованными ортогональными весами. Ортогонализация и параметризация весов свёрточных ядер использовалась для повышения эффективности интерполяционного алгоритма машинного обучения, которое в свою очередь повлекло за собой повышение эффективности метода компрессии в целом. Вычислительные эксперименты продемонстрировали выигрыш предложенного алгоритма на реальных видеосигналах.

#### БЛАГОДАРНОСТИ

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

#### ЛИТЕРАТУРА

- [1] Свириденко, В.А. Высокоэффективная компрессия видео: новые подходы / В.А. Свириденко // Цифровая инфраструктура для трансформации экономики: задачи и возможности. – 2020. – С. 61-66.
- [2] Петров, Д.А. Проблемы обучения нейронных сетей на основе объемных данных видеоконтента / Д.А.Петров, Р.И.Салаватов // Дифференциальные уравнения и смежные проблемы. – 2018. – С. 255-258.
- [3] Gashnikov, M.V. General Structure of an Machine Learning Method for Compression of Images / M.V. Gashnikov // VIII International Conference on Information Technology and Nanotechnology (ITNT). – 2022. – P. 1-4. DOI: 10.1109/ITNT55410.2022.9848719.
- [4] Dynamic Scenes Data Set [Electronic resource]. — Access mode: <http://vision.eecs.yorku.ca/research/dynamic-scenes> (request date 19.10.2022).