

Разработка общей структуры метода компрессии изображений, основанного на машинном обучении

М.В. Гашников

Самарский национальный исследовательский университет им. академика С.П. Королева

Самара, Россия

mih-fastt@yandex.ru

Аннотация—Эта статья предлагает архитектуру основанного на машинном обучении (МО) метода для компрессии изображений и двумерных сигналов. Метод использует основанное на квадродереве представление двумерного сигнала. Это представление позволяет использовать характерные особенности различных этажей этого квадродерева для настройки МО-методов. Это представление также позволяет нам предсказывать отсчёты двумерного сигнала на основе менее детальных этажей квадродерева и кодировать отклонения предсказания. Мы используем МО-методы на всех этапах метода компрессии. В частности, мы используем основанные на МО алгоритмы суперразрешения и встраивания цифровых двумерных сигналов, для предсказания отсчётов сигнала: свёрточные и генеративно-состязательные нейросети, а также автоэнкодеры. Мы проводим вычислительные эксперименты по исследованию эффективности основанного на МО предсказания на наиболее детальном этапе основанного на квадродереве представления двумерного сигнала при компрессии. Мы экспериментально доказываем, что исследуемый основанный на МО предсказатель отсчётов имеет существенное преимущество на реальных цифровых двумерных сигналах.

Ключевые слова— суперразрешение, предсказание, двумерный сигнал, нейросети, архитектура метода.

1. ВВЕДЕНИЕ

Научная значимость и актуальность исследований в области сжатия цифровых изображений [1] с использованием методов машинного обучения (в частности, методов глубокого обучения) [2] обусловлена целым рядом факторов: ростом пространственного разрешения, качества цветопередачи и в целом качества восприятия визуальных данных, ростом объема хранимых и передаваемых визуальных данных в сетях общего пользования, ростом ожиданий пользователей относительно качества получаемой ими видеoinформации и т.д.

В данной работе разрабатывается структура основанного на машинном обучении (МО) метода компрессии изображений. Метод использует безыбыточный вариант основанного на квадродереве представления изображения [3], позволяющий учитывать при настройке МО структурные особенности изображения, проявляющие на различных масштабных уровнях. Такое представление позволяет интерполировать отсчёты изображения по более прореженным версиям того же самого изображения с последующим энтропийным кодированием постинтерполяционных остатков. Для интерполяции

используются алгоритмы [2, 4] суперразрешения и доопределения цифровых изображений, основанные на машинном обучении: свёрточные (CNN) и генеративно-состязательные нейросети (GAN), а также автоэнкодеры. Использование МО позволяет существенно повысить эффективность отдельных этапов метода компрессии, что, в конечном итоге, приводит к существенному повышению эффективности всей технологии в целом.

2. МЕТОД КОМПРЕССИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Предложенная структура основанного на МО метода компрессии изображений показана на рис. 1.

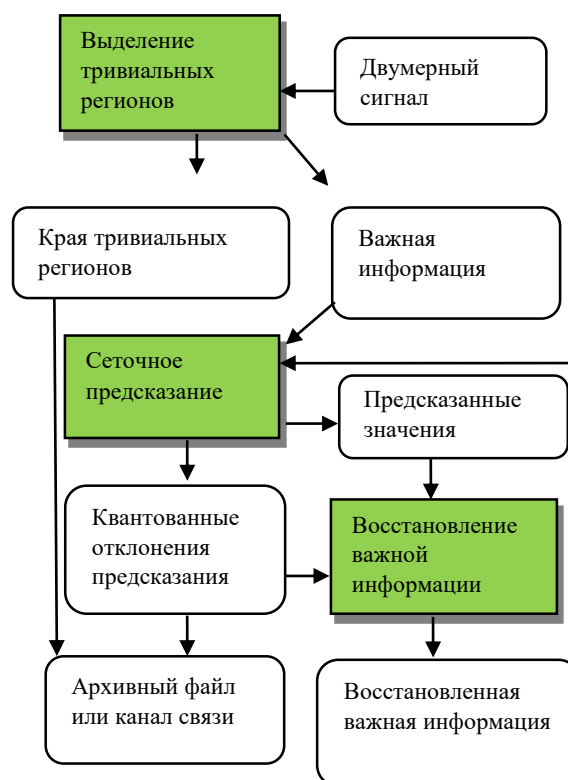


Рис. 1. Архитектура основанного на машинном обучении метода компрессии двумерных сигналов

Мы выделяем тривиальные регионы в двумерном сигнале. Мы используем основанную на МО сегментацию для этого. Тривиальные регионы (например, внутренняя часть озёр, морей, рек, лесов в снимках Земной поверхности и т.п.) не содержат важной информации. Мы помещаем в архивный файл или канал связи только края тривиальных регионов, так как мы можем использовать алгоритмы встраивания

изображений для заполнения этих регионов при декомпрессии. Мы архивируем только «важную информацию», которая расположена вне этих регионов.

Мы используем основанный на квадродереве алгоритм [3] для компрессии регионов, содержащих «важную информацию». Мы представляем двумерный сигнал внутри «важных» регионов как квадродерево. Этажи квадродерева представляют из себя непересекающиеся части прореженного двумерного сигнала внутри «важных» регионов.

Мы компрессируем и сразу декомпрессируем этажи квадродерева последовательно, начиная с наименее детального этажа. Мы получаем искаженные (декомпрессированные) значения отсчётов в результате обработки каждого этажа квадродерева внутри регионов двумерного сигнала, содержащих важную информацию.

Мы предсказываем отсчёты каждого этажа квадродерева на основе отсчётов менее детальных этажей посредством нейросетей. Мы используем декомпрессированные значения этих отсчётов в качестве базы для этого предсказания.

Используемые свёрточные нейросети модифицируются посредством ортогонализации и параметризации настраиваемых свёрточных ядер в целях повышения устойчивости и сокращения времени обучения. Ортогонализация выполняется через внедрение специализированных штрафов на линейную зависимость свёрточных ядер в функцию потерь при обучении нейросети. Для параметризации свёрточных ядер применяется подход, заимствованный из методологии построения мультипараметрических вейвлетов.

3. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

Экспериментальный раздел настоящей работы посвящён исследованию эффективности нейросетевой интерполяции в рамках предлагаемого метода компрессии на натуральных изображениях (см. рис. 2).

В качестве нейросетевого интерполятора использован метод [5]. На рис. 3 показаны результаты интерполяции на наиболее детальном уровне основанного на квадродереве представления изображения при компрессии. Видно, что исследуемый нейросетевой интерполятор имеет существенное преимущество над билинейным интерполятором.

4. ЗАКЛЮЧЕНИЕ

Предложена архитектура основанного на МО метода компрессии изображений. Метод использует алгоритмы сверхразрешения, встраивания и кластеризации, основанные на CNN и GAN. Вычислительные эксперименты подтвердили существенное преимущество исследованных алгоритмов на реальных изображениях.

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда (проект 22-21-00662).



Рис. 2. Пример исходного натурального тестового изображения

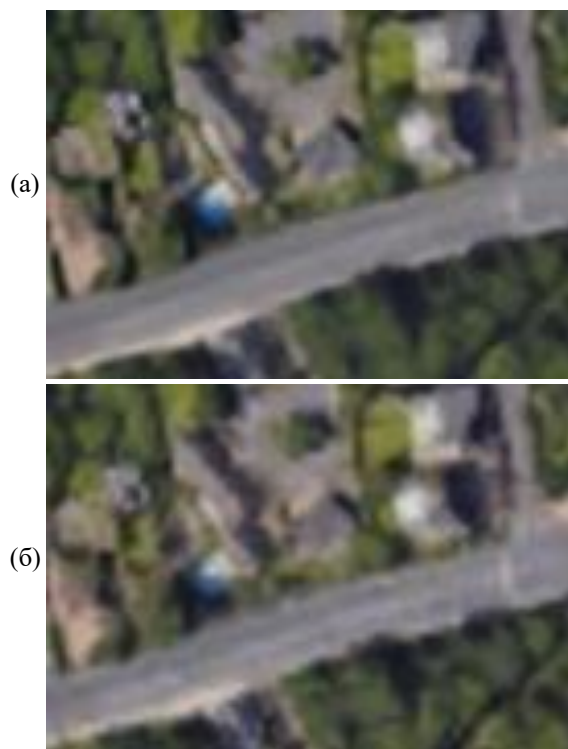


Рис. 3. Нейросетевая (а) и билинейная (б) интерполяция на наиболее детальном этапе основанного на квадродереве представления изображения при компрессии

ЛИТЕРАТУРА

- [1] Зевадова, Л.А. Краткий обзор методов и алгоритмов компрессии зашумлённых изображений / Л.А. Зевадова, Ю.С. Бехтин // Современные научные исследования: теория, методология, практика. – 2018. – Т. 1. – С. 26-35. DOI: 10.26425/2658-3445-2018-1-26-35.
- [2] Коротеев, М.В. Обзор некоторых современных тенденций в технологии машинного обучения / М.В. Коротеев // E-Management. – 2018. – Т. 1, № 1. – С. 26-35. DOI: 10.26425/2658-3445-2018-1-26-35
- [3] Сергеев, В.В. Информационная технология компрессии изображений в системах оперативного дистанционного зондирования / М.В. Гашиков, Н.И. Глузов, В.В. Сергеев // Известия Самарского научного центра РАН. – 1999. – № 1. – С. 99-107.
- [4] Гашиков, М.В. Параметризованная интерполяция при комплексировании многомерных сигналов различного разрешения / М.В. Гашиков // Компьютерная оптика. – 2020. – Т. 44, № 3. – С. 436-440. DOI: 10.18287/2412-6179-CO-696.
- [5] Wang, X. Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data / X. Wang, L. Xie, C. Dong, Y. Shan // IEEE/CVF International Conference on Computer Vision. – 2021. – P. 1905-1914. DOI: 10.48550/arXiv.2107.10833.