

АЛГОРИТМ УЛУЧШЕНИЯ КАЧЕСТВА МЕДИАФАЙЛОВ

Минько Д. А.

Факультет информационных технологий и управления, Белорусский государственный университет
информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: 98_minko@mail.ru

Рассматриваются основные подходы по улучшению качества медиафайлов. Обозревается подход на основе свёрточной нейронной сети, а также метод разреженного кодирования

ВВЕДЕНИЕ

Изначально многие медиафайлы сжаты с целью уменьшить размер файлов, что делает их идеальными для обмена и размещения в интернете. Вследствие этого, при попытке увеличить или использовать изображение повторно, оно может выглядеть зернистым или заретушированным. Повысить качество медиафайлов можно путем корректировки внешних атрибутов, цвета и контрастности изображения при помощи редактора изображений.

I. ОСНОВНЫЕ ПОДХОДЫ

Известно несколько основных подходов по улучшению изображения: предсказательные модели, краевые методы, статистические методы, методы основанные на паттернах, метод преобразования гистограмм, подход на основе свёрточной нейронной сети.

1. Предсказательное моделирование основывается на построении, управлении и расчете моделей при помощи техник аппроксимации;
2. Краевые методы - методы о нахождении решения заданного дифференциального уравнения (системы дифференциальных уравнений), удовлетворяющего краевым (граничным) условиям в концах интервала или на границе области;
3. Статистический метод – метод, основанный на анализе большой выборки данных (что и является статистикой);
4. Методы основанные на паттернах – методы, в основе которых лежат готовые паттерны (шаблоны) для решения конкретных задач.
5. Метод преобразования гистограмм. В данном методе пытаются достичь равномерности распределения яркостей обработанного изображения.
6. Свёрточная нейронная сеть - нейронная сеть, использующая множество идентичных копий одного и того же нейрона. Это позволяет сети иметь ограниченное число параметров при вычислении больших моделей. Свёрточная нейронная сеть, однажды обучив нейрон, использует его во множе-

стве мест, что облегчает обучение модели и минимизирует ошибки.

Наилучшее качество дают метод преобразования гистограмм и подход на основе свёрточной нейронной сети.

II. ОБЗОР ПОДХОДА НА ОСНОВЕ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Пусть у нас есть изображение Y , которое мы хотим восстановить ($F(Y)$), и оно должно быть максимально похожее на основное изображение высокого разрешения X . Данный подход состоит из трех операций:

1. Извлечение и представление патчей: эта операция извлекает (перекрывающиеся) патчи из изображения Y с низким разрешением и представляет каждый патч как вектор большой размерности.
2. Нелинейное отображение: эта операция нелинейно отображает каждый многомерный вектор на другой многомерный вектор. Каждый сопоставимый вектор концептуально представляет патч высокого разрешения. Эти векторы составляют другой набор карт характеристик.
3. Реконструкция: эта операция объединяет вышеприведенные патч-представления высокого разрешения для генерации окончательного изображения высокого разрешения. Ожидается, что это изображение будет похоже на X .

Первая операция выражается как:

$$F_1(Y) = \max(0, W_1 * Y + B_1), \quad (1)$$

Где W_1 и B_1 представляют фильтры и смещения соответственно, а «*» обозначает операцию свертки. На выходе после первой операции мы получим для каждого патча n_1 -мерный массив. Во второй операции происходит преобразование этого массива в n_2 -мерный, увеличивая размер первого массива для повышения точности. Каждый из выходных n_2 -мерных векторов концептуально представляет собой патч высокого разрешения, который будет использоваться для реконструкции:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2). \quad (2)$$

В традиционных методах предсказанные перекрывающиеся участки с высоким разрешением часто усредняются для получения окончательного полного изображения. Усреднение можно рассматривать как предопределенный фильтр на наборе карт характеристик (где каждая позиция является «сплюсненной» векторной формой патча высокого разрешения). Основываясь на этом, мы определяем сверточный слой для получения окончательного изображения с высоким разрешением:

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3. \quad (3)$$

Хоть вышеописанные три операции интуитивно разные, но они все сводятся к той же форме, что и сверточный слой. Объединив их вместе, образуется сверточная нейронная сеть (рисунок 1). В этой модели все веса фильтрации должны быть оптимизированы.

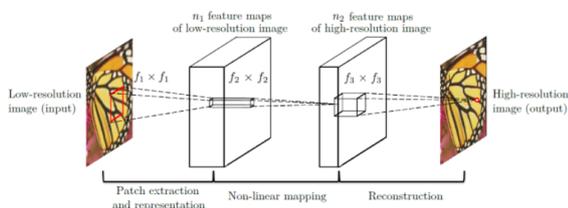


Рис. 1 – Визуализация работы сверточной нейронной сети

III. МЕТОД РАЗРЕЖЕННОГО КОДИРОВАНИЯ

Сверточная нейронная сеть имеет связь с методом разреженного кодирования (рисунок 2). Разреженное кодирование – это метод обучения представления, цель которого – поиск разреженного представления входных данных в форме линейных комбинаций базовых элементов, вместе с самими базовыми элементами. В методе разреженного кодирования предположим, что патч с низким разрешением $f_1 \times f_1$ извлекается из входного изображения. Затем патч проектируется в словарь (с низким разрешением). Если размер словаря равен n_1 , то это эквивалентно применению n_1 линейных фильтров на входное изображение. Это показано на левой части рисунка 3. Затем программа разреженного кодирования будет итеративно обрабатывать n_1 коэффициентов. Выходы этого решения – n_2 коэффициентов. Как правило, $n_1 = n_2$ в случае разреженного кодирования. Коэффициенты n_2 и представляют патч высокого разрешения. В этом случае программа разреженного кодирования ведёт себя как частный случай оператора нелинейного отображения. Это изображено на средней части рисунка 3. Вышеупомянутые коэффициен-

ты n_2 затем проектируются в другой словарь (с высоким разрешением) для создания фрагмента высокого разрешения. Перекрывающиеся участки усредняются, это эквивалентно линейным свёрткам на n_2 карты характеристик. На выходе получаем патчи с высоким разрешением размера $f_3 \times f_3$, тогда линейные фильтры имеют эквивалентную пространственную поддержку размера $f_3 \times f_3$.

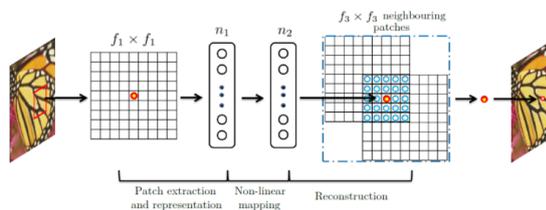


Рис. 2 – Связь сверточной нейронной сети с методом разреженного кодирования

Вышеупомянутая последовательность показывает, что метод на основе разреженного кодирования можно рассматривать как своего рода сверточную нейронную сеть (с другим нелинейным отображением). В данной сверточной сети словарь с низким разрешением, словарь с высоким разрешением. Нелинейное отображение вместе со средним вычитаемым усреднением задействованы в оптимизируемых фильтрах. Таким образом, этот метод оптимизирует сквозное отображение, состоящее из всех операций. Ниже приведена таблица, в которой показаны результирующие значения качества изображения и времени срабатывания, в зависимости от числа фильтров.

$n_1 = 128$ $n_2 = 64$		$n_1 = 64$ $n_2 = 32$		$n_1 = 32$ $n_2 = 16$	
PSNR	Time (sec)	PSNR	Time (sec)	PSNR	Time (sec)
32.60	0.60	32.52	0.18	32.26	0.05

IV. ВЫВОД

Улучшение качества медиафайлов – довольно важный элемент повседневной жизни и сферы технологий, который применяется буквально везде, где есть цифровое изображение. Данное направление ежедневно дорабатывается многими крупными компаниями, такими как Google, Apple. В результате проделанной работы было установлено, что подход на основе сверточной нейронной сети является одним из самых практичных и востребованных. Как минимум в сфере распознавания объектов или улучшения качества медиафайлов, или, проще говоря, везде, где необходимо обучать нейронную сеть. Такие результаты достигаются благодаря операциям свёртки.